Stable isotope labelling in the analysis of bacterial antigen processing

Victoria J. Thomas

Ph.D.

The University of Edinburgh, 2007

I declare that I have written this thesis and that the work is my own. This body of work has not been submitted for any other degree or professional qualification.

Miss Victoria J. Thomas

Table of Contents

List of Figures & Tables

List of abbreviations

α3 (IV)	α3 antigen chain of type IV collagen
HAc	Acetic acid
ACN	Acetonitrile
DTT	Dithiothreitol
EBV	Epstein Barr virus
ESI	Electrospray ionisation
HPLC	High performance liquid chromatography
IEX	Ion exchange
IPTG	Isopropyl-beta-D-thiogalactopyranoside
LB	Luria broth
MALDI	Matrix assisted laser desorption ionisation
MS	Mass spectrometry
MS/MS	Mass spectrometry / mass spectrometry
nLC	Nano-flow Liquid chromatography
OMP	Outer membrane protein
1D	One dimensional
RP	Reverse phase
TFA	Tri-fluoroacetic acid
TOF	Time of flight
Tris	Tris-(hydroxymethyl) methylamine
Tris-HCl	Tris-(hydroxymethyl) methylamine hydrochloride
TTCF	Tetanus toxin C fragment
2D	Two dimensional
MBP	Myelin basic protein
HEL	Hen egg lysozyme

List Of Equipment	
37°C Orbital incubator	· Gallenkamp
CO2 cell incubator	· Fisher
Centrifuges	· Sorvall OTD, Sorvall RC5B, Eppendorf 5810R
Sonicator	· MSE Soniprep 150
Fluorescent ELISA plate reader	· Cytofluor 4000® Series Fluorescence Multi-well Plate Reader, Applied Biosystems
Mass spectrometer	· Voyager™ DE-PRO Biospectrometry ™ Workstation (PerSeptive Biosystems) · 400 well Teflon coated MALDI mass spectrometry plate
High performance liquid chromatography system	Items supplied by Michrom BioResources, Inc. · MAGIC 2002 HPLC system · MAGIC HPLC Xcaliber Control System · MAGIC Temperature Controlled Autosampler · Magic Bullet strong IEX column Items supplied by LC packings, USA · C18 RP Column C18 (15cm x 75µm) · C18 Peptide Captrap
Speed vac	· Savant AES 1000 / AES 1010
Ball bearing homogeniser	· EMBL, Germany
Western blotting tank	· BioRad

List of Chemicals, Reagents, Growth Media & Buffers

List of Chemicals & Reagents

Detergents	· 7% w/v sodium N lauroylsarcosinate ["sarkosyl" solution] · 0.02% CHAPS Both detergents purchased from Sigma
Mass spectrometry reagents	· Carbon-13 glucose (Cambridge Isotope laboratories) · α-cyano-4-hydroxycinnamic acid (Sigma) · Peptide Standards Kit, calibration mixture 2, (PerSeptive Biosystems)
Chromatography reagents	HPLC super purity solvents and acids All purchased from ROMIL: · Pure water · Acetonitrile · Methanol · Tri-fluoroacetic acid (TFA) Other chromatography reagents · KCl (Sigma) · Imidazole (Sigma) · NaCl (Sigma)
Enzymes & substrates	· Modified porcine trypsin (Promega) · Horse radish peroxidase conjugated anti human IgG (Sigma) · 4-methylumbelliferyl n-acetyl β-D-glucosamide (Sigma)

Protein Assay Reagents	· 12.5% w/v aqueous sodium carbonate · 0.1% w/v aqueous copper sulphate · Folin & Ciocaltues phenol reagent (supplied by BDH) · 2g/L bovine serum albumin

List of Growth Media

Luria Broth (LB) medium (100mls)	All reagents used to make bacterial growth media purchased from Sigma · Bacto-tryptone (1g) · Yeast extract (0.5g) · NaCl (0.5g)
5 X M9, pH 7.0 (500mls)	· 0.24M Na2HPO4 (32g) · 0.11M KH2PO4 (7.5g) · 0.043 M NaCl (1.25g)
Salt mix (10mls)	· 4mM $ZnSO_4$ · 1mM $MnSO_4$ · 0.7mM H3BO3 · 0.7mM CuS04
M9 Minimal Media (1L)	· 400mls H20 · 200 mls 5*M9 · 1ml salt mix · 2mls 1M MgSO4 · 2mls 50Mm CaCl2 · 400µl 5mM FeCl3 · 800µl thiamine (0.5mg/ml) · 9.4% NH4Cl2 · 15mls Glucose (20%) · 1ml carbenicillin (50mg/ml)

List of Buffers

Bacterial outer membrane extraction buffers	· 0.05M sodium phosphate buffer pH 7.4 (0.05M $NaH_2PO_4.2H_20$ and 0.05M Na_2HPO_4; pH 7.4) containing 0.15M sodium chloride
SDS- PAGE buffers	· **40% w/v aqueous acrylamide solution** acrylamide supplied by BDH Electran · **15g/l w/v ammonium persulphate solution** ammonium persulphate supplied by BDH analar · **Separating gel buffer (Double strength)** 0.75 M Tris/HCl pH 8.8, 0.2% SDS · **Stacking gel buffer (Double strength)** 0.25M Tris/HCl pH 6.8, 0.2% SDS · **Sample buffer (Double strength)** 0.125 M Tris / HCl pH 6.8, 4% SDS glycerol, 2% 2-mercaptoethanol, 0.002% bromophenol blue · **Electrode buffer** 0.025M Tris, 0.192 M glycine, 0.1% SDS, pH 8.3
Tris-tricine buffers	All Tricine buffers supplied by Invitrogen · Novex® Tricine SDS Running Buffer (10x) · Novex® Tricine SDS Sample Buffer (2x)
Coomassie blue staining buffers	· **Solution 1** (1g Coomassie brilliant blue R-250, 500ml propanol-2-ol, 200ml acetic acid, 1300ml distilled water) · **Solution 2** (100mg Coomassie brilliant blue R-250, 200ml propanol-2-ol, 200ml acetic acid, 1600ml distilled water) · **Solution 3** (48mg Coomassie brilliant blue R-250, 200ml acetic acid, 1800ml distilled water) · **Solution 4** (800ml methanol, 200ml acetic acid, 1000ml distilled water)

	· **Solution 5** (200ml acetic acid, 1800ml distilled water)
LPS silver staining buffers	· **LPS fixative** (25% propan-2-ol, 7% acetic acid) · **Oxidising solution** (0.7% periodic acid / 25% propan-2-ol) · **Ammoniacal silver nitrate** (0.36% NaOH, 19.4% silver nitrate) · **Developer** (0.005% citric acid in 0.019% formaldehyde solution)
Immunoblotting buffers	· **Western blotting electrode buffer** 0.025M Tris / 0.192M glycine pH 8.3, 20% methanol · **Tris buffered saline (TBS), pH 7.5** (0.02M Tris / HCl, 0.5M NaCl) · **Tween tris buffered saline (TTBS), pH 7.5** (0.02M Tris / HCl, 0.5M NaCl, 0.025% tween 20) · **Blocking solution** 3% gelatin (BioRad) in TBS · **Antibody buffer** 1% gelatine in TBS · **Horse radish peroxidase colour development solution solution** **A** – 60mg 4-chloro-1-naphthol (Bio-Rad HRP colour reagent) & 20ml methanol (BDH Analar) **B** - 60µl hydrogen peroxide (BDH Analar, 30 % w/v H_2O) & 100mlTBS Freshly prepare just before use. mix A and B together and add to nitrocellulose
Lysosomal extraction & digestion buffers	·**Homogenisation buffer** (10mM triethanolamine, 10mM acetic acid, 1mM Na_2 EDTA, 0.25M 60% sucrose [w/v])

	· **27% Percoll solution** (10mM triethanolamine, 10mM HAc, 1mM Na_2 EDTA, 0.25M 60% sucrose, 27% Percoll) · **50mM citrate buffer, pH 5.5**
Protein identification buffers	· **Reducing buffer** 10mM DTT in 50-100µl of 100mM $NH_4(HCO_3)_2$ / 5% ACN · **Alkylating Buffer** 50mM iodoacetamide in 50-100µl 100mM $NH_4(HCO_3)_2$ · **Digestion buffer** 50mM tris-phosphate buffer (pH 7.4)
Mass spectrometry buffers	**Sample buffer** · 0.1% TFA in purified water Both TFA and water purchased from ROMIL
Chromatography buffers	**Reverse phase chromatography buffers** · **Buffer A:** 3% Acetonitrile · **Buffer B:** 90% Acetonitrile, 0.08% TFA · **Sample buffer (same as "buffer A")** **Cation Exchange chromatography buffers used in the development of HPLC system (Chapter 4)** · **Buffer A**: 10% ACN, 0.08% TFA · **NH4OH salt plugs:** 50mM, 100mM, 250mM, 500mM, 1M **Cation Exchange chromatography buffers used in the 2D separation of lysosomal extracts (Chapter 5)** · **Buffer A:** H_20 + 0.5% v/v acetic acid · **Buffer B:** H_20 + 0.5% v/v acetic acid + 35% ACN + 500mM KCl

	Affinity chromatography buffers (His-Tag) · **Conditioning buffer:** 25mM imidazole, 500mM NaCl, 500mM Tris-Cl · **Eluting buffer:** 500 mM imidazole, 500mM NaCl, 500mM Tris-Cl

Acknowledgements

I would like to thank my supervisors, Professor Ian R. Poxton and Dr Richard G. Phelps for their help and support over the last four years. I am also very grateful to Professor Colin Watts (Dundee University) for providing me with invaluable materials for my research project. I would also like to thank the medical school for providing me with my PhD scholarship. Lastly, I would like to acknowledge all the members in my research group for their encouragement during my PhD.

Abstract

The mechanism by which antigens are processed and presented to T cells exerts a powerful influence on all immune responses, including those directed against pathogens. Knowledge of the peptides produced by antigen presenting cells for presentation to T cells is important in understanding why individuals have different abilities to combat infection and may prove critical in the design of new vaccines. The aim of this project was to attempt to use a novel metabolic labelling technique to enable biochemical identification of potential T cell epitopes in bacterial antigens. The technique is based upon labelling antigens with the stable isotope carbon-13, and subsequent identification of antigen-derived peptides using mass spectrometric isotopic composition analysis. The ultimate goal of the work was to pulse antigen presenting cells with labelled bacteria and identify MHC class II-associated bacteria-derived peptides with the objective of identifying potential vaccine candidates.

Towards this goal, outer membrane proteins were prepared and analysed by polyacrylamide gel electrophoresis and immunoblotted with serum from healthy individuals. The results showed that the serum antibodies bound to a number of outer membrane proteins, confirming that they contained B cell epitopes. As a first step to demonstrate that the outer membrane proteins also contained T cell epitopes, it was shown that whole bacteria and bacterial antigens can be successfully labelled with carbon-13 such that all derived peptide fragments may be distinguished by isotopic analysis. However, due to the complexity of the peptide mixture generated by a tryptic digest of the outer membrane proteins, it was critical to combine isotope analysis with excellent chromatographic resolution in order to identify carbon-13 labelled peptides. To address this problem, a new high sensitivity 2D nanoflow liquid chromatography system was developed and its capabilities compared with conventional 1D reverse phase chromatographic techniques. Two dimensional separation prior to mass spectral analysis substantially improved the separation of complex biological mixtures leading to an improvement in the identification of carbon-13 labelled peptides in such mixtures.

Two dimensional chromatography and carbon-13 labelling was then applied to investigate the processing of tetanus toxin C fragment (TTCF) by lysosomes isolated from EBV transformed B lymphoblastoid cell lines. Whilst large TTCF

fragments were demonstrably generated during processing, small (<3kd) peptides suitable for binding to MHC class II molecules were extremely scarce. Further experiments suggested that this was unlikely to be a consequence of insufficient sensitivity of the carbon-13 technique, but probably a consequence of rapid destruction of the smaller fragments by lysosomal enzymes and high stability of the larger digestion fragments of TTCF.

It was concluded from these studies that the metabolic labelling of bacterial proteins with carbon-13 is applicable to the analysis of processed products of bacterial antigens in complex peptide mixtures. Small peptides, with sizes in the range bound to MHC class II molecules, appear to be very sparse in processing compartments, presumably because of rapid proteolysis. Peptides bound to MHC class II are known to be relatively resistant to proteolysis so in future work it is proposed to examine TTCF-derived peptides bound to MHC class II molecules rather than free within lysosomes.

1. Introduction

T lymphocytes are cells of the immune system that orchestrate the cell-mediated arm of the acquired immune response. This includes the protective immune responses generated against pathogens and cancer cells, and the disease causing responses that cause autoimmune disease. Both $CD4^+$ and $CD8^+$ T cells recognise protein antigens in the form of 'processed' antigen derived peptides 'presented' on the surface of antigen presenting cells bound to MHC molecules (Unanue 1984; Townsend and Bodmer 1989). The activation of $CD4^+$ and $CD8^+$ T cells is the result of the two separate pathways within cells (figure 1.1) known as the MHC class I and MHC class II processing pathways (Pamer and Cresswell 1998; Pieters 2000). Activation of T cells by peptides presented by the relevant pathway leads to the most appropriate immune response for clearance of viral and bacterial pathogens.

$CD8^+$ T cells recognise peptides bound to MHC class I molecules which are constitutively expressed at the surface of the majority of nucleated cells. This may reflect the fact that every cell within the body is susceptible to infection. Peptides feeding into this pathway are derived from cytosolic proteins that have been digested in the cytoplasm by a structure called the proteasome. These peptides are subsequently transported into the endoplasmic reticulum (ER) by the transporter associated with antigen processing (TAP) and loaded onto class I molecules for display at the cell surface to $CD8^+$ T cells. During infection, viral proteins and those derived from intracellular bacteria (e.g. *Listeria monocytogenes*) feed into this pathway (reviewed in Pamer and Cresswell 1998).

$CD4^+$ T cells recognise peptides bound to MHC class II molecules that are constitutively expressed on the surface of dendritic cells, macrophages, and B cells. These cells also express the appropriate co-stimulation required to activate T cells and are thus termed 'professional antigen presenting cells'. During infection, class II expression can also be induced upon other cell types such as epithelial cells by γ-interferon. Thus, a wider range of cells has the ability to act as antigen presenting cells during an immune response. Peptides feeding into this pathway include those derived from exogenous proteins internalised into the cell. These proteins are digested in the endosomal–lysosomal pathway where they are loaded onto MHC

class II molecules within specialised MHC class II loading compartments (MIIC) for expression at the cell surface to $CD4^+$ T cells. A minor pathway also exists in which recycling MHC class II molecules bind to their ligands in early endosomes (reviewed in Pieters 2000).

The MHC class II pathway is expected to be of more importance in the processing and presentation of extracellular bacterial antigens. For this reason, the following section will take a more in depth look at antigen processing and presentation within the MHC class II pathway. However, it should be noted that MHC class I molecules can also present peptides derived from exogenous antigens to $CD8^+$ T cells. This alternative processing pathway for exogenous antigens will be discussed in section 1.1.8.

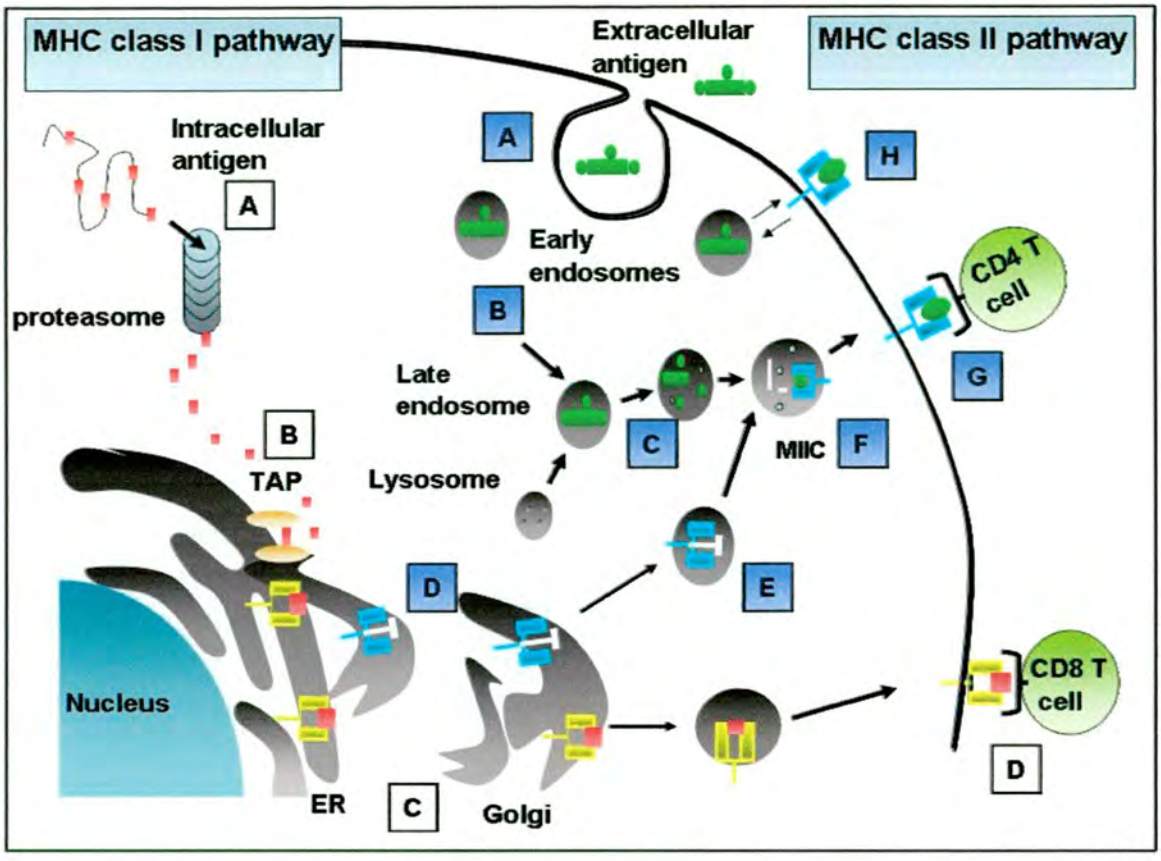


Figure 1.1 MHC class I and class II processing pathways
MHC class I pathway: A) Intracellular antigen is digested into peptides by a cytosolic structure called the proteasome. **B)** Peptides released by the proteasome are transported into the ER by TAP where they bind to MHC class I molecules. **C)** MHC class I – peptide complexes travel through the Golgi body and are directed to the cell surface. **D)** Surface MHC class I – peptide complexes present peptide to $CD8^+$ T cells.
MHC class II pathway: A) Extracellular antigen is engulfed into the cell by a form of endocytosis. **B)** Early endosomes mature into late endosomes, and eventually fuse with lysosomes **C)** Antigen is digested into peptides by proteolytic enzymes **D)** MHC class II heterodimers are synthesised in the ER and associate with the invariant chain to form nonameric complexes ($\alpha\beta$Ii)3. **E)** These complexes pass through the Golgi body and are directed to the compartment for peptide loading (MIIC) **F)** In the MIIC, the invariant chain is removed from the MHC class II groove in exchange for peptide. **G)** MHC class II – peptide complexes are transported to the cell surface where they present peptide to $CD4^+$ T cells. **H)** Recycling MHC class II molecules from the cell surface may also bind to antigen in early endosomes.

1.1 Antigen processing and presentation in the MHC class II pathway

The mechanisms involved in MHC class II presentation of extracellular antigens are usefully considered as two distinct but overlapping pathways. First is the pathway taken by extracellular antigen, which includes the key stages of antigen uptake into the antigen presenting cells and the processing of antigen into peptides suitable for MHC class II binding. In this pathway, protein antigens are internalised into the cell by a variety of mechanisms and are subsequently delivered to the endosomal / lysosomal network for enzymatic degradation. In antigen presenting cells, some antigen derived peptides may be rescued from proteolytic destruction by incorporation into MHC class II / peptide complexes, which are subsequently transported to the cell surface for $CD4^+$ T cell recognition. The second pathway is that taken by MHC class II molecules from their site of synthesis to the relevant MHC class II loading compartments. These pathways overlap at several places providing opportunities for class II molecules to bind antigen-derived peptides. The following sections will initially follow the pathway taken by antigen engulfed into antigen presenting cells and will then focus upon the pathway taken by class II molecules from their site of synthesis to the site of peptide loading within antigen presenting cells.

1.1.1 Antigen uptake

Antigen uptake is critical for the processing and presentation of the majority of antigens. The rate at which antigen is internalised is an important determinant of the efficiency of presentation, with antigen presenting cells being able to utilise several uptake mechanisms which differ greatly in their specificity, efficiency and capacity (Lanzavecchia 1996). These uptake mechanisms include both non-specific (pinocytosis and macropinocytosis) and specific (receptor mediated endocytosis and phagocytosis) mechanisms. The predominant mechanisms of uptake employed by an antigen presenting cell influences their contribution towards the development of an immune response and also the functional outcome.

1.1.1.1 Fluid phase uptake

Fluid phase antigens that are unable to bind to cell surface receptors may be internalised non-specifically along with fluid captured by the cell during pinocytosis (Swanson and Watts 1995). During this process, invaginations of the cell membrane form, pinch off as pre-endosomal vesicles and rapidly fuse with early endosomes. The formation of clathrin coated pits commonly precedes pinocytosis, forming clathrin coated vesicles of between 85-110 nm in diameter. Macropinocytosis is a distinct mechanism that is best characterised by marked membrane ruffling followed by the envacuolation of extracellular fluid within large (0.5-5 μm) uncoated vesicles. Whilst both macrophages and endothelial cells are capable of macropinocytosis upon stimulation, immature DC are the only cells that constitutively carry out macropinocytosis (West *et al.*, 2000). Macropinocytosis is thought to be an important mechanism by which bacteria and larger sized particles (>0.5 μm) are engulfed into immature dendritic cells before they migrate to secondary lymph nodes to generate an immune response (Xiang *et al.*, 2006).

B cells are very poor at fluid phase uptake, being estimated to pinocytose extracellular antigen at a 100-1000 fold lower level than macrophages (Chesnut *et al.*, 1982; Lanzavecchia 1990). As pinocytosis represents the sole mechanism of uptake for non-specific B cells, these B cells are very inefficient at presenting antigen to T cells during primary immune responses.

1.1.1.2 Receptor-mediated endocytosis

Both macrophages and dendritic cells express a variety of membrane receptors that can bind to antigen (e.g. mannose receptor) or to the Fc portion of circulating immunoglobulins (e.g. FcγR). Like pinocytosis, receptor mediated endocytosis is commonly characterised by the formation of clathrin coated pits followed by the formation of pre-endosomal vesicles (Edeling *et al.*, 2006). However, receptor mediated endocytosis has clear quantitative and qualitative advantages in comparison to non-specific fluid uptake mechanisms. For example, antigen specific B cells present antigen to T cells at 100-1000 fold lower antigen concentrations compared to non-specific B cells (Rock *et al.*, 1984; Lanzavecchia 1985). This is thought to be the result of a concentrating effect upon the surface of

the antigen presenting cell, leading to faster and more efficient antigen uptake of much lower levels of soluble antigen (Lanzavecchia 1990). Similarly, mannose receptor-mediated uptake of antigens results in a approximately 100-fold more efficient presentation to T cells, as compared to antigens internalized via fluid phase (Engering *et al.*, 1997). Further evidence to support enhanced immunogenicity following receptor mediated mechanisms has been provided by the conjugation of antigens usually internalised by fluid phase pinocytosis to ligands for cell surface receptors. For example, antigen conjugates targeting surface immunoglobulin produce equivalent in-vitro T cell responses at 500-1000 fold lower antigen concentrations and are more immunogenic in vivo (Lakey *et al.*, 1988; Snider and Segal 1989; Barnes and Mitchell 1995). Similar enhancement has been reported for conjugates targeting other cell surface receptors including the transferrin receptor (McCoy *et al.*, 1993), Fcγ receptors (Gosselin *et al.*, 1992), the α2 macroglobulin mannose receptors (Chu and Pizzo 1993), and the mannose receptor (Sallusto *et al.* 1995).

In addition to enhancing the immune response, receptor mediated endocytosis can also influence the functional outcome of an immune response by influencing epitope production. One proposed mechanism for this is by the preferential targeting of antigen to different compartments. For example, there is now evidence to suggest that different components of the B cell receptor complex may target internalised antigens to different processing compartments. The B cell receptor complex consists of a surface immunoglobulin (sIg) which is associated with an Igα / Igβ heterodimer. It has been shown that the cytoplasmic domains of Igα and Igβ contain signalling domains that selectively target ligand – receptor complexes for presentation on newly synthesised or recycling MHC class II molecules respectively (Bonnerot *et al.*, 1995). Further experiments demonstrated that the signalling motifs in Igα and Igβ targeted antigens to distinct sub-cellular organelles, leading to the production of different T cell epitopes (Amigorena and Bonnerot 1998). Similarly, it has been demonstrated that different Fc receptors lead to the production of different T cell epitopes from the same antigen that may be the result of antigen delivery to different processing compartments by different Fc receptors (Amigorena *et al.* 1998; Amigorena and Bonnerot 1999). In addition, the mannose receptor recycles

constitutively between the plasma membrane and early endosomes where it releases its ligand (Stahl *et al.*, 1980). Thus, the delivery of antigen to different processing compartments by different receptors or different components of the same receptor may influence the epitopes produced and thus the outcome of an immune response.

A second mechanism by which receptors may influence the immune response is by directly interfering with the processing of antigens. This has been demonstrated most clearly for the B cell receptor. Interference of the B cell receptor following internalisation of receptor-ligand complexes 'sculpts' antigen processing (Davidson and Watts 1989; Watts *et al.*, 1998; Antoniou and Watts 2002). This form of epitope directed processing has been shown to influence the production of T cell epitopes from tetanus toxin C fragment (TTCF) in a number of B cell lines. The bound immunoglobulin imparts a 'footprint' upon the antigen, protecting areas from enzymatic processing and MHC class II binding. Uptake of radiolabelled TTCF into a number of B cell lines has revealed that different immunoglobulin receptors 'footprinted' different areas of the antigen, leading to the production of different processing fragments detected using SDS-PAGE. Antibody guided processing has also been shown to suppress the loading of certain T cell epitopes located within the footprinted area whilst boosting the presentation of others (Simitsek *et al.*, 1995).

1.1.1.3 Phagocytosis

Phagocytosis is a ligand-induced form of endocytosis by which large particles such as microbes and tissue debris are engulfed into antigen presenting cells (Brown 1995; Silverstein 1995; Aderem and Underhill 1999). In addition to antigen presenting cell, other cell types are capable of phagocytosis such as epithelial cells of the thyroid, bladder and retina (Zeligs and Wollman 1977; de Boer *et al.*, 1996; Sarangarajan and Apte 2005; Strauss 2005) However, it is only antigen presenting cells that express a large enough number of receptors to allow them to phagocytose a wide variety of antigens. For example, they possess receptors for complement components, LPS in bacterial cell walls, and mannans in yeast wall cells. This allows them to phagocytose wide range of microbial antigens. Following the binding of antigen to phagocytes, actin rearrangement leads to the formation of membranous

extensions called 'pseudopods' which engulf large particles (>0.5 μm) into the cell, targeting the antigen for destruction in the endocytic pathway.

Both macrophages and immature dendritic cells readily engulf large particulate antigens by phagocytosis. For instance, macrophages are able to internalise up to 50% of their cell surface area in one round of phagocytosis (Steinman *et al.* 1983). It has been recently shown that a large amount of the phagosomal membrane is of ER origin, suggesting that the ER and plasma membranes may fuse together during phagosome formation to compensate for the vast amounts of membrane used during phagocytosis (Gagnon *et al.*, 2002; Watts 2002; Edeling *et al.*, 2006). In contrast to macrophages and dendritic cells, B cells do not carry out phagocytosis.

Following internalisation, phagosomes undergo maturation by fusing with endosomes and lysosomes, to form a phagolysosome that contains a variety of potent enzymes required for antigen degradation and epitope production (Aderem and Underhill 1999). There is evidence to suggest that epitopes released following phagocytosis may be different to those released by non-specific uptake mechanisms. Von Delwig *et al.*, (2002) demonstrated that phagocytosis of viable *Streptococcus pyogenes* by macrophages leads to the production of an epitope from the surface M5 protein ($M5_{17-31}$) that is not produced following uptake by macropinocytosis. Similarly, uptake by macropinocytosis results in the production of a different epitope from the surface M5 protein ($M5_{308-319}$) that is not produced following phagocytosis. Other results from the group demonstrated that these two epitopes have different processing requirements and that $M5_{17-31}$ was bound by mature and recycling MHC class II molecules whereas $M5_{308-319}$ was bound by newly synthesised MHC class II molecules (Delvig and Robinson 1998; Delvig and Robinson 1998). It was thus postulated that *Streptococci pyogenes* is delivered to different processing compartments following uptake by these two mechanisms, leading to the generation of different epitopes. Therefore, the type of uptake mechanism employed by a cell influences the production of epitopes from an antigen. This may lead to the generation of a more diverse repertoire of T cell epitopes, thus influencing the functional outcome of an immune response.

1.1.2 Antigen processing

Once antigens have been engulfed into antigen presenting cells, they are degraded into peptide fragments by proteolytic enzymes for MHC class II binding and subsequent display to T cells (Riese *et al.*, 1996; Riese and Chapman 2000; Turk *et al.*, 2000; Villadangos and Ploegh 2000; Watts 2001). The acidic environment of the endocytic pathway is important for the majority of proteolytic enzymes to work at their optimum, with a decrease in pH from 6.5 to 4.5, being paralleled by increased enzyme activity and harsher proteolysis.

The decrease in pH also contributes to antigen denaturation and unfolding which is important in allowing the access of enzymes to their cut sites within the antigen (Jensen 1993). A second important factor influencing denaturation is the reduction of disulfide bonds within the antigen. Disulfide bond reduction has been shown to influence the production of epitopes from several antigens and thus T cell recognition (Collins *et al.*, 1991; Haque *et al.*, 2001; Maric *et al.*, 2001; Li *et al.*, 2002). Oxidation of protein antigens is also thought to contribute to antigen unfolding and to the exposure of epitopes recognised by specific T cells (Carrasco-Marin *et al.*, 1998).

The traditional view of antigen processing is that enzymes work in a redundant manner to degrade antigens, with many enzymes being able to carry out the same role during degradation. To support this idea, analysis of MHC class II associated peptides revealed that there are no common cleavage sites at the termini of eluted peptides, indicating that any of the proteases may act if given the opportunity (Rudensky *et al.* 1991). In addition, it has been shown in-vivo that mice lacking several of the enzymes previously thought to be important in antigen processing are not deficient in antigen processing (Villadangos *et al.*, 1997; Deussing *et al.*, 1998). However, there is now increasing evidence to suggest that individual enzymes can play a more important and non-indispensable role in the processing of antigens. The roles of these enzymes together with others implicated in antigen processing will be discussed in further detail in the following sections.

1.1.2.1 Cathepsins

Both cathepsins D and B were originally suggested to play an important role in antigen processing. This was due to in-vitro studies implicating both of these enzymes in antigen processing together with the high expression of these enzymes in antigen presenting cells (Watts 1997). However, it was later revealed that knockout mice deficient in either of these two enzymes were not deficient in the presentation of T cell epitopes from antigens, dismissing their earlier role (Villadangos *et al.*, 1997; Deussing *et al.*, 1998). However, additional studies have once again implicated these proteases in antigen processing. For example, Driessen *et al.*, (2001) demonstrated that cathepsin B was a key enzyme in the in-vivo degradation of an immune complex following Fc γ R-mediated uptake in antigen-presenting cells. In addition, Zhang *et al.*, (2000) demonstrated that the inhibition of cathepsin B in mice immunized with ovalbumin, modulated the specific immune response from a T helper 2 (T_H2) type towards a T_H1 response. It has also been demonstrated that the processing of exogenous glutamate decarboxylase (GAD) is sensitive to inhibition of cathepsin D (Lich *et al.*, 2000). Thus, cathepsin B and D almost certainly have an important role, but are not essential, possibly because other endosomal proteases have similar specificity.

Both cathepsins S and L have been shown to play a role in antigen processing. A recent study analysing the MHC peptide repertoire of embryonic fibroblast lines deficient in either cathepsin S or L reported that epitope generation for a subset of antigens is specifically regulated by the activity of both of these enzymes (Hsieh *et al.*, 2002). Another in-vitro study demonstrated that cathepsin S dominated the processing of myelin basic protein following digestion with lysosoml extracts isolated from human B lymphoblastoid cells (Beck *et al.*, 2001). Several in-vivo studies have also shown a role for cathepsin S in antigen processing. Using cathepsin deficient mice, it was demonstrated that cathepsin S plays an independant role in the degradation of immune receptor complexes delivered via Fcγ R into antigen presenting cells. In a second study, antigen presenting cells failed to present specific T cell epitopes from hen egg lysosome in cathepsin S deficient mice (Pluger *et al.*, 2002). Most recently, it was demonstrated that intestinal epithelial cells present

a class II-bound endogenous peptide to naive $CD4^+$ T cells *in vivo* in a Cat S-dependent fashion (Beers *et al.*, 2005).

Recently, another member of the cathepsin family has been shown to play a dominant role in antigen processing. The serine protease cathepsin G plays a critical role in the in-vitro processing of myelin basic protein (MBP) following digestion with lysosomes isolated from primary B cells, (Burster *et al.*, 2004). Unlike other cathepsins, cathepsin G is not endogenously synthesised by human B lymphocytes. Rather, it is internalised into lysosomes from the plasma membrane where it subsequently participates in antigen processing. Due to its cell surface expression, it has been suggested that cathepsin G may also play a role in extracellular antigen degradation. However, this remains to be elucidated.

1.1.2.2 Asparaginyl endopeptidase

Asparagine endopeptidase (AEP) is a novel cysteine protease that has strict specificity, cleaving only after accessible asparagine residues in antigens. AEP activity was first described during experiments showing that AEP is critical in the initiation of processing of the model antigen, tetanus toxin C fragment (TTCF) (Manoury *et al.*, 1998). The dominant role of AEP in the digestion of TTCF was the first example of an enzyme playing a critical non-redundant role in antigen processing. Using lysosomal extracts isolated from lymphoblastoid B cells, AEP was shown to be the dominant enzyme in the digestion of TTCF, cleaving after 3 asparagine residues to generate 5 processing products detectable by SDS-PAGE analysis. Suppression of AEP activity in human B cells was found to inhibit presentation of TTCF to a number of T cell clones, whilst the pre-processing of TTCF with AEP accelerated presentation. Further experiments demonstrated that mutagenesis of the 3 AEP cleavage sites in TTCF together with inhibition of AEP activity slowed down the kinetics of antigen presentation as well as the subsequent presentation of a number of T cell epitopes (Antoniou *et al.*, 2000). Cleavage by AEP after 1 or 2 asparagine residues in TTCF was responsible for this influence upon antigen presentation.

As a result of all of these experiments, a model was postulated for the digestion of TTCF (figure 1.2). In this model, it is hypothesised that an initial

cleavage by AEP acts to 'unlock' TTCF. Together with denaturation, unlocking of the antigen in this manner allows other proteases as well as AEP to carry out further cleavages at hydrophobic sites previously inaccessible to proteolysis. Although AEP is critical in the unlocking step in TTCF digestion, different enzymes may perform this initiating step for different antigens. It should be noted, however, that it is not clear whether all or even most antigens require unlocking prior to digestion. It has been shown that in AEP deficient mice, presentation to primary T cells of OVA and myelin oligodendrocyte glycoprotein, two antigens that contain asparagine residues within or in proximity to the relevant epitopes was unimpaired (Maehr *et al.*, 2005). Thus, the importance of AEP in the production of T cell epitopes from a wide range of antigens remains to be elucidated.

In contrast to its action upon TTCF, it has also been shown that AEP action can be destructive. Studies using the autoantigen myelin basic protein revealed that AEP activity destroyed a T cell epitope, preventing its display on class II MHC molecules (Manoury *et al.*, 2002; Watts *et al.*, 2003). This is due to the presence of an accessible AEP processing site at Asn 94 within the epitope. It has been shown during in-vitro experiments that the presentation of this self epitope inversely correlates with the level of AEP. As a consequence, it has been suggested that destruction of this epitope within the thymus by AEP may prevent the establishment of T cell tolerance towards this 'cryptic' epitope. Later changes in antigen processing may result in the presentation of this peptide, leading to the generation of auto-reactive T cells. For instance, it has been suggested that natural deaminidation of AEP cut sites in 'ageing' self antigens (e.g. MBP) may lead to the presentation of previously cryptic epitopes to T cells and thus autoimmune disease (Moss *et al.*, 2005).

1.1.2.3 γ-interferon-inducible lysosomal reductase (GILT)

GILT is an enzyme that is thought to play an important role in disulphide bond reduction within lysosomes (Arunachalam *et al.* 2000). Initial experiments demonstrated that GILT is optimally active between pH 4-5 and is capable of catalysing the disulphide bond reduction of F(ab)'2 and intact immunoglobulin during both in-vitro and in-vivo experiments. It was therefore hypothesised that

GILT may aid in the unfolding of protein antigens in the acidic endocytic environment. Further experiments demonstrated that disulphide bridge reduction catalysed by GILT is important in the production of T cell epitopes from protein antigens (Maric *et al.*, 2001). In GILT free mice, the presentation of two major epitopes from hen egg lysozyme (HEL) was found to be partially or completely abrogated, directly implicating this enzyme in disulphide bond cleavage and the generation of certain T cell epitopes. These results clearly show a non-redundant and critical role for GILT in the reduction of disulphide bonds in at least some protein antigens. As a consequence, it has been hypothesised that antigen unfolding in the presence of GILT aids in the access of both the initiating protease and secondary proteases to their cut sites (Watts 2001). This would explain the importance of GILT in the generation of some T cell epitopes from antigens.

Although GILT plays a key role in the processing of antigens, there may also be other yet unidentified enzymes involved in endosomal redox reactions. For example, it has been shown that the action of GILT is not responsible for the reduction of intrachain disulfide bonds required for the presentation of a H2-E^d-restricted epitope in the influenza hemagglutinin antigen (Sinnathamby *et al.*, 2004). Antigen presenting cells (fibroblasts) obtained from GILT-free mice presented the H2-E^d-restricted epitope, and introduction of GILT had no impact on presentation. The authors concluded that there might be mechanisms for GILT-independent reduction in the late endosome.

1.1.2.4 Other enzymes

There is limited information on which other enzymes are thought to play key roles in the processing of antigens. However, it has been demonstrated that a mutation in HEL to create a dibasic (lysine-arginine) motif adjacent to a subdominant determinant (HEL23-32) resulted in a ~10 fold increase in presentation of the determinant (Schneider *et al.*, 2000). The identity of the enzyme that made this cleavage is unknown although it was suggested to be a member of the pro-protein convertases (Seidah and Chretien 1997). In addition, the use of serine and metalloprotease inhibitors have implicated both these classes of enzymes in early endosomal processing of antigens and in the generation of certain T cell epitopes

from antigens (Delvig and Robinson 1998; Musson *et al.*, 2003). Furthermore, the inhibition of serine proteases has been shown to block the presentation of a number of epitopes that also require further lysosomal processing before presentation. Thus, initial processing in early endosomes may be essential prior to further processing within lysosomes in the production of certain T cell epitopes.

1.1.2.5 MHC guided processing

It has been traditionally assumed that peptides released during antigenic degradation are rescued from proteolytic destruction by binding to MHC molecules in specialised endocytic compartments. However, how do peptides survive in this harsh proteolytic environment long enough for MHC class II molecules to bind? It has been postulated that this may be because MHC class II molecules protect T cell epitopes from complete degradation by binding to partially unfolded antigen (Sercarz *et al.*, 1993; Sercarz and Maverakis 2003). Following the binding of MHC II molecules, endopeptidases can then act at other processing sites within the antigen and the bound peptides ultimately trimmed by exopeptidases (e.g. cathepsin B, cathepsin H).

There is now increasing evidence to support that MHC guided processing does play a role in the 'sculpting' of antigens during proteolysis. It has been demonstrated that MHC class II molecules can bind to flexible regions in intact antigens as well as full-length reduced antigens (Lee *et al.*, 1988; Sette *et al.*, 1989). In addition, it has been shown that only limited processing of TTCF is required by AEP before MHC class II binding can occur. Therefore, it is possible that MHC class II molecules may bind to antigen in endocytic compartments after very little processing. In support of this, studies have revealed that larger peptides (3000-7000 Da) can be eluted from I-A or I-E murine class II molecules if care is taken (Castellino *et al.*, 1998). Furthermore, the presence of a long hen egg lysozyme derived polypeptide was detected bound to two different MHC class molecules ($I\text{-}A^k$ and $I\text{-}E^k$) within the endocytic pathway following the pulsing of antigen presenting cells with HEL (Germain and Margulies 1993).

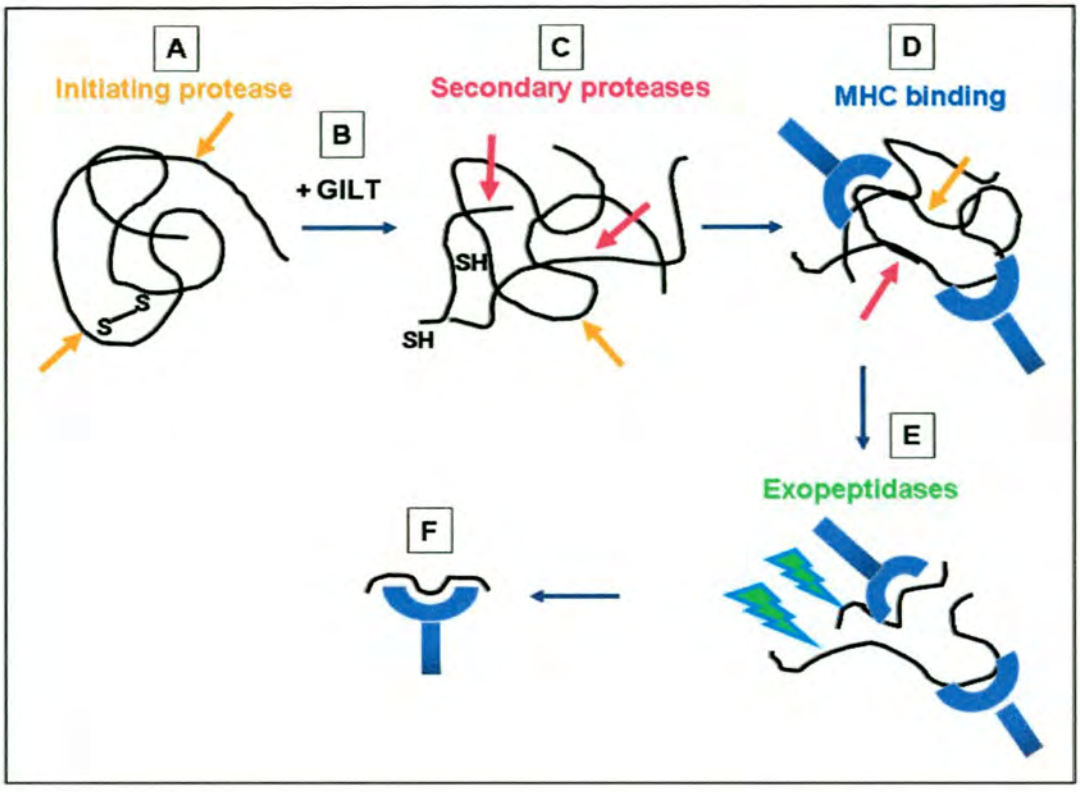


Figure 1.2 Proposed model for digestion of antigens in the endocytic environment

(A) An initiating protease (e.g. AEP) makes the initial cleavage(s) which acts to unlock the protein **(B)** Reduction of disulfide bonds by enzymes such as GILT aids in antigen unfolding and thus proteolytic digestion **(C)** Endopeptidases (e.g. cathepsin D) act together with the initiating protease to further digest the antigen at previously hidden hydrophobic sites **(D)** MHC class II molecules may protect the destruction of epitopes by binding to regions of the partially unfolded antigen **(E)** Exopeptidases (e.g. cathepsin B) trim the MHC class II bound peptides to lengths eluted from the cell surface shown in **(F)**.

1.1.3 Structure and function of MHC class II molecules

The structure of MHC molecules, and more specifically the peptide binding groove, is critical in understanding the basis of peptide binding. The following section will look at structural differences between MHC class I and II molecules and how these differences influence the peptides that can bind. It will also briefly discuss the structural basis of peptide binding and how polygeny and polymorphism in the MHC leads to different MHC molecules presenting different sets of peptides to T cells.

1.1.3.1 Genetic organisation of the MHC

The MHC is located on chromosome 6 in humans and on chromosome 17 in mice. In humans, it extends over ~4 centimorgans of DNA, consisting of 4 $x10^6$ base pairs and more than 200 genes (Beck and Trowsdale 2000). Figure 1.3 shows the general organisation of the MHC class I and class II genes in both humans and in mice. The organisation of the MHC complex is similar in both humans and mice, with genes being organised into class I and class II regions. In humans, the MHC is termed the Human Leukocyte Antigen (HLA), comprising of genes encoding both class I molecules called HLA-A, B, C and class II molecules called HLA-DP, DQ and DR. In mice, the MHC is called the Histocompatibility-2 (H-2) complex. Like in humans, there are 3 class I genes called H2-K, -D, and –L, and 3 class II genes called H2-A and –E. Within the MHC class II genes, there are genes encoding DO and DM molecules that are important in the MHC class II processing pathway and LMP and TAP molecules that are important in the MHC class I processing pathway. In both humans and mice, the MHC complex also contains class III genes encoding a number of proteins important in the immune response including complement components and cytokines.

1.1.3.1.1 Polygeny and polymorphism within the MHC

The MHC is polygenic with several genes encoding both class I and class II molecules. In humans, there are three class I α-chain genes (HLA-A, -B, and –C) and three pairs of MHC class II α and β chain genes (HLA-DR, DP, and DQ). In many cases, the HLA-DR cluster contains an extra β gene whose product can pair with the

DR α chain. This means that the three sets of genes give rise to four types of MHC class II molecule. All of the MHC molecules encoded by these genes have different peptide specificities, allowing them to bind to different sets of peptides. As a result of polygeny, every person will express at least 3 different MHC class I molecules and 3 different MHC class II molecules at the surface of their antigen presenting cells. Thus, polygeny increases the chance that peptides generated from a single antigen will be bound by an MHC molecule and presented to T cells. This is important for the generation of immune responses against infectious agents on an individual level.

With the exception of DRA1, the class II loci are also highly polymorphic, with there being over 200 alleles of some human MHC class I and class II genes (Robinson *et al.*, 2003; Schreuder *et al.*, 2005). As a result of the polymorphic nature of the MHC, most individuals are heterozygous at MHC loci, inheriting different alleles from their parents. Since MHC expression is co-dominant, individuals may have up to 6 different MHC class I molecules and 8 different MHC class II molecules. This increases the diversity already available through polygeny at the individual level. The extensive polymorphism at each locus is also protective at the community level, with different individuals possessing different MHC molecules and therefore having differing susceptibility to infection.

Polymorphisms are confined to codons that determine 2 classes of residues in MHC molecules. The first class of residues are those flanking the pockets within the peptide binding groove (Babbitt *et al.*, 1985; Babbitt *et al.*, 2005). These residues are important in peptide binding because they determine which peptides can bind to the MHC molecule. The second class of residues contribute to parts of the α helices which contribute to the surface engaged by T cell receptors (Zinkernagel and Doherty, 1997). Allelic variation thus results in class II molecules with very similar structures but with different peptide specificities and TCR binding characteristics.

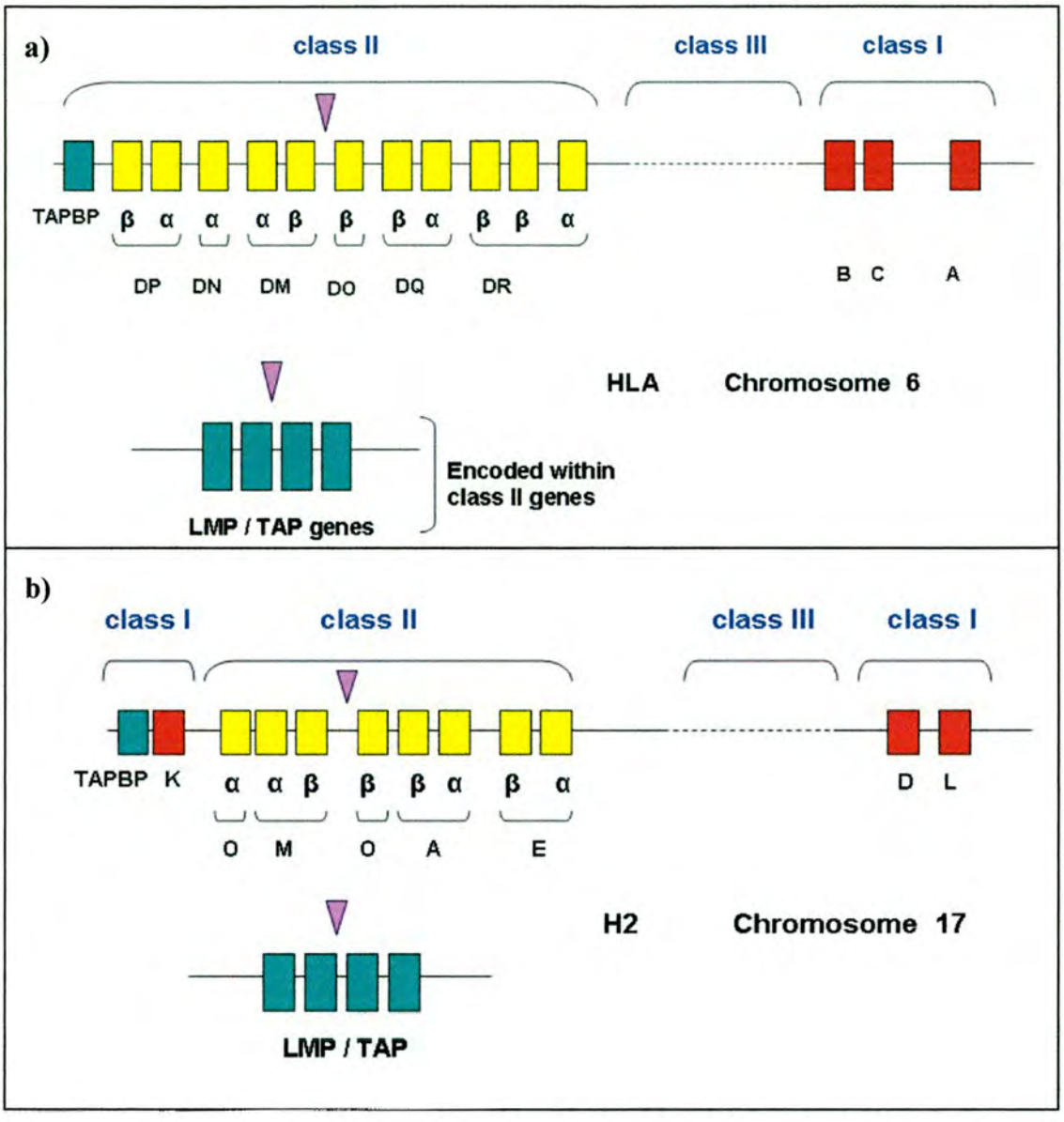


Figure 1.3 The Major Histocompatibility Complex (MHC) in humans and mice

The organisation is shown for the principal MHC genes in both **(a)** humans and **(b)** mice. The arrangement of MHC genes is similar in both species, with clusters of MHC class I, class II, and class III genes being observed. However, in mice, the MHC class I gene H-2K has translocated relative to the human MHC so that the class I region in mice is split into two.

1.1.3.2 Structure of MHC molecules

MHC molecules are heterodimeric transmembrane glycoproteins that are expressed at the cell surface. The structures of class I and class II molecules are shown in figure 1.4. Crystallographic studies have shown that class I and class II molecules have remarkably similar 3-dimensional structures (Bjorkman *et al.*, 1987; Brown *et al.*, 1993). Both have an antigen binding groove with a beta pleated sheet floor bordered on 2 sides by alpha helices. This arrangement allows peptides to be securely bound with elements of their structure visible to T cells. Despite this broad similarity, there are important differences between class I and class II molecules which profoundly influence the peptides they bind.

MHC class I molecules comprise a single transmembrane glycoprotein non-covalently associated with β2 microglobulin. The class I peptide binding groove is closed at both ends limiting the sizes of peptides that can stably bind. Interactions between the amino and carboxyl termini of bound peptides with residues in pockets at each end of the groove are important for stable binding. Peptides shorter than the length of the groove (~9 amino acids) bind poorly; peptides longer than this (up to 12 residues) may be accommodated only if the peptide backbone can bulge up out of the middle of the groove.

MHC class II molecules consist of two non-covalently linked membrane glycoproteins called the alpha (α) and beta (β) chains. Both of these are encoded in the MHC complex and consist of a short cytoplasmic domain and transmembrane portion, followed by a membrane proximal domain called the α2 and β2 domains respectively. These are both similar in amino acid sequence and structure to immunoglobulin constant regions. The membrane proximal domain in each chain is followed by a membrane distal domain called the α1 or β1 domain, which together form the peptide-binding groove. Unlike MHC class I molecules, the class II peptide binding groove is open at both ends, and as a consequence, can accommodate long peptides that protrude at either side of the binding cleft (Demotz *et al.*, 1989; Chicz *et al.*, 1993; Rammensee *et al.*, 1993; Falk *et al.*, 1994). Thus, antigen presenting cells can present a large amount of peptides bound to only a small set of different class II molecules.

Structural analysis of the 3-dimensional structure of class II - peptide complexes has revealed the structural basis of peptide binding (Stern *et al.*, 1994; Ghosh *et al.*, 1995). Peptides are bound with a 9 amino acid sequence clasped in the groove. Longer peptides protrude at either or at both ends of the groove. The bound segment of a peptide is often called the 'core binding sequence' because of its major importance in the class II / peptide interaction. Peptide / class II interactions, most important in stabilising the bound peptide, occur between peptide backbone atoms and conserved residues / main chain atoms in parts of the class II molecules flanking the peptide binding groove. Thus, binding is stabilised mainly by interactions that are independent of the bound peptide's sequence.

These binding-favourable interactions depend on the bound peptide adopting a particular conformation in the groove that requires the peptide's side chain residues to be accommodated within pockets in the walls of the peptide-binding groove. The crystal structures obtained for DR1 and DR3 with bound peptide show a very similar pattern of pockets arranged along the walls and base of the peptide-binding groove. However, the class II amino acid residues that flank the pockets in the DR1 and DR3 structures are quite different. These residues determine the character of the pockets (charge, hydrophobicity and size) and hence the ease with which different peptide side chain residues can be accommodated, placing different constraints on the sequences of peptides that can be stably bound.

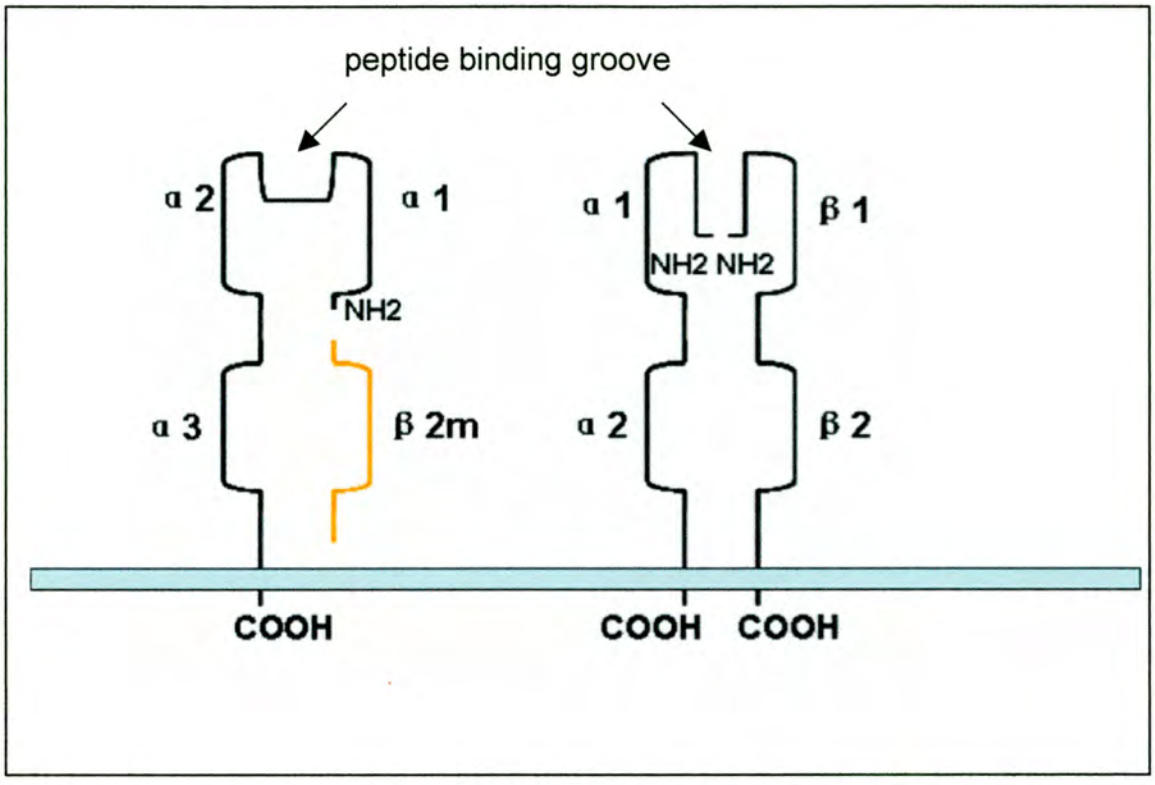


Figure 1.4 Structure of MHC class I and class II molecules

MHC class I and class II proteins are heterodimeric transmembrane glycoproteins. MHC class I molecules consist of an α chain encoded within the MHC which is non covalently associated with a β2M molecule encoded outside of the MHC. The α chain folds into 3 domains, with the α1 and α2 domains forming the binding site and the α3 domain containing a transmembrane domain. In contrast, MHC class II molecules consist of an α and β chain both encoded within the MHC. The α and β chain fold into 2 domains each, with the α1 and β1 domains forming the peptide binding groove and the α2 and β2 domains containing transmembrane anchors.

1.1.4 MHC class II biosynthesis and trafficking

In order for peptide binding to occur, MHC class II molecules must enter peptide-containing compartments of the endocytic pathway. MHC class II molecules are synthesised in the ER and subsequently targeted to class II loading compartments for peptide binding. To prevent MHC class II molecules from binding to peptides before their arrival at the relevant class II loading compartments, newly synthesised MHC class II molecules associate with a type II transmembrane glycoprotein called the invariant chain (Ii) to form nonameric complexes, (αβIi)3 (Jones *et al.*, 1979; Cresswell 1994). This association prevents cytosolic peptides in the ER from binding to MHC II molecules (Roche and Cresswell 1990; Teyton *et al.*, 1990; Busch *et al.* 1996). A signalling motif present in the cytoplasmic domain of the invariant chain acts to target the class II complexes to the appropriate endocytic compartments for class II binding (Bakke and Dobberstein 1990; Lotteau *et al.*, 1990; Pieters *et al.*, 1993). These MHC class II - peptide complexes are subsequently transported to the cell surface for presentation to T cells. Surface class II / peptide complexes are stable with a turn over rate of many days, but continuously recycle through early endosomes and a proportion also reaches later endocytic compartments. Thus, class II molecules are detectable throughout the endocytic system as well as in the ER, the Golgi and at the cell surface. Identifying the precise route by which class II traffics between these compartments, and in particular the locations at which peptides are bound, is still not complete.

The compartment to which newly synthesised MHC class II – invariant chain complexes are directed may not be identical in all cell types, but in B cells and dendritic cells most nascent MHC class II – invariant chain complexes rapidly enter a distinct compartment called the MHC class II compartment or MIIC (Harding 1995). The MIIC has ultra-structural, enzymatic, and electrophoretic properties distinct from compartments of the classical endocytic pathway. It has a density similar to late endosomes but has a distinct multivesicular morphology on electron microscopy and lacks late endosomal (mannsoe-6-phosphate receptor) and early endosomal (transferrin receptor) markers (Peters *et al.*, 1991). It also contains molecules which are important for the formation of class II / peptide complexes. More recently, class

II loading compartments (CIIV) have been identified as MHC class II rich compartments in B cells (Amigorena *et al.*, 1994). These differ from MIIC in that they structurally resemble early endosomes (Pierre and Mellman, 1998). In B cells, peptide loading can occur in either MIIC or CIIC (Bryant and Ploegh 2004). It is thought that most nascent MHC molecules reach these endocytic compartments without visiting the cell surface but the exact route is uncertain. Some MHC class II – invariant chain complexes are detectable at the cell surface and rapidly recycle into endosomes (Roche *et al.* 1993). Recycling MHC class II – invariant chain complexes and class II / peptide complexes presumably supply the populations of class II in other endocytic compartments, some of which also contain HLA-DM (Guagliardi *et al.*, 1990).

1.1.5 Sites of MHC class II maturation and peptide loading

Class II molecules are widely distributed among the intracellular compartments of antigen presenting cells and can probably bind to peptides at several locations, the most important depending upon the antigen and cell type (Castellino and Germain 1995). At least for some antigens, the majority of antigen-derived peptide / class II molecules are composed of newly synthesised class II molecules, so the way nascent class II become loaded with peptide in MIIC has been closely studied in several cell types (Harding 1996).

Following delivery of MHC class II - Ii complexes to the MIIC, the Ii is degraded by proteases to leave the class II associated peptide (CLIP) within the class II peptide binding cleft (Blum and Cresswell 1988; Pieters *et al.*, 1991; Chapman 1998). During the maturation of MHC class II molecules, the Ii chain goes through the intermediate forms lip 22 (MW 22 KDa) and lip 10 (MW 10 KDa) before its conversion into CLIP. The enzyme AEP has been shown to be one of the enzymes that can perform the initial conversion of Ii into lip 22 (Manoury *et al.*, 2003). Cathepsin S plays a critical role in the terminal processing of lip-10 into CLIP in B cells, dendritic cells, and intestinal epithelial cells (Riese *et al.*, 1996; Nakagawa *et al.*, 1999; Shi *et al.*, 1999; Beers *et al.*, 2005). Cathepsin L has the same function as cathepsin S in thymic epithelial cells (Nakagawa *et al.*, 1998).

The removal of CLIP in exchange for peptide is promoted by the chaperone HLA-DM (Denzin and Cresswell 1995; Sherman *et al.* 1995; Sloan *et al.*, 1995). The MIIC is the compartment with the highest concentration of HLA-DM. However, it has now been demonstrated that HLA-DM is located throughout the endocytic pathway. In addition to its critical function in exchanging CLIP for peptide, DM also functions to catalyse the release of low-stability peptides, allowing high affinity peptides to ultimately form complexes with MHC class II molecules (Katz *et al.*, 1996; Weber *et al.*, 1996; Kropshofer *et al.* 1997). In B cells, DM is regulated by the chaperone HLA-DO (Liljedahl *et al.*, 1996; Denzin *et al.*, 1997). DO is thought to inhibit the activity of DM in all but the most acidic late endosomal compartments, to which the majority of BCR- internalised antigens are targeted. Thus, DO might promote the presentation of peptides derived from BCR-internalised antigens (Alfonso *et al.*, 2003; Brocke *et al.*, 2003; Bryant and Ploegh, 2004).

1.1.6 Alternative sites of MHC class II peptide loading

An alternative minor pathway has also emerged, in which mature MHC class II molecules recycled from the cell surface bind to antigen in early endosomes (Griffin *et al.*, 1997; Pinet and Long 1998; Robinson and Delvig 2002). The binding of epitopes to recycled MHC class II molecules is proving critical for the presentation of a growing number of epitopes, including those derived from HEL, influenza virus haemagglutinin, myelin basic protein, and tetanus toxin. Recycled MHC class II molecules may bind to flexible regions within intact antigens or alternatively following partial unfolding and degradation. It is highly probable that the epitopes released in early endosomes may be destroyed by harsher proteolysis in later processing compartments. Indeed, it has been demonstrated that peptides released by early endosomes from antigens differ from those released in later processing compartments (Pinet *et al.*, 1994). Thus, the release of different epitopes in early and late endosomes may give rise to a varied epitope repertoire for T cell recognition (Watts 1997; Robinson and Delvig 2002). The dependence of peptide loading in early endosomes upon HLA-DM is unknown, with several epitopes being shown to be loaded onto recycled class II molecules in a HLA-DM independent manner.

In addition to peptide loading within late and early endosomes, there is now evidence to suggest that MHC class II loading can also occur within the phagosomal pathway, with class II molecules, invariant chain and HLA-DM all being found in phagosomes (Ramachandra *et al.*, 1999). It has been shown that phagosomal compartments contain MHC class II molecules and that purified phagosomes containing antigen coupled latex beads can activate specific T cell hybridomas (Ramachandra *et al.*, 1999; Ramachandra and Harding 2000). Further experiments carried out by the group demonstrated that phagosomal processing and presentation of antigen-coupled latex beads was dependant primarily on nascent MHC class II molecules delivered from intracellular sites (e.g. endocytic compartments). Other groups, however, have demonstrated that the binding of recycled MHC molecules to peptides in phagosomes is essential for the presentation of a number of epitopes (Pinet *et al.*, 1994; Cella *et al.*, 1997; Balaji and Boom 1998; Delvig and Robinson 1998; Musson *et al.*, 2002). Despite these results, it is not clear whether the bulk of peptides released from particulate antigens, bind to MHC class II molecule present within phagosomes, or are transported into endosomal compartements for MHC loading (Ramachandra *et al.*, 1999).

1.1.7 MHC Class II associated peptides

Information about the properties of peptides bound to MHC molecules came from structural determination of the MHC molecules and by the characterisation of eluted peptides (Rotzschke *et al.*, 1990; Van Bleek and Nathenson 1990; Fremont *et al.* 1992; Rammensee *et al.*, 1993). As predicted by crystallographic studies, peptides eluted from class I molecules were approximately 9 amino acid residues in length (Bjorkman *et al.*, 1987; Schumacher *et al.*, 1991; Rammensee *et al.*, 1993). The same techniques later applied to MHC class II elution studies revealed that much longer peptides of between 15-25 amino acids in length bind to MHC class II molecules, consistent with the crystal structure (Demotz *et al.*, 1989; Rudensky *et al.*, 1991; Chicz *et al.*, 1992; Hunt *et al.*, 1992; Riberdy *et al.*, 1992; Chicz *et al.*, 1993; Newcomb and Cresswell 1993).

In class II elution studies, peptides were derived mostly from self membrane-associated and endosomal proteins including MHC molecules, invariant chain and

proteolytic enzymes. A minority derived from exogenous proteins. These are mainly proteins present in the extracellular fluid at high concentrations (several mg/ml) such as albumin and haemoglobin, and proteins for which there are very efficient uptake mechanisms such as apolipoprotein B-100 and transferrin. Analysis of MHC class II peptides revealed that MHC class II eluted peptides typically consisted of peptides of heterogeneous length derived from the same region of the protein. These sets of peptides were termed 'nested sets' and were found to overlap over at least 12 core binding residues but differ in length at both the N and C termini, reflecting the structure of the MHC class II binding groove. Despite the variable lengths of class II peptides, residues in a 9 amino acid core were shown to be critical for either MHC binding or TCR contact (Stern *et al.*, 1994; Rammensee 1995; Johansen *et al.*, 1996).

Structural features of peptides facilitating peptide binding to specific MHC molecules became clear following sequence characterisation of MHC eluted peptides and co-crystallisation of MHC molecules bound to peptide (Garrett *et al.* 1989; Fremont *et al.* 1992; Matsumura *et al.*, 1992). This led to the identification of peptide motifs and anchor residues within peptides eluted from specific class II molecules (Falk *et al.*, 1991; Falk *et al.*, 1994). Peptide binding motif is the term given to the common amino acids shared by peptides binding to particular class I and II molecules with high affinity. They are often termed 'anchor residues' because they stabilise binding of the peptide inside the groove. Crystallographic studies have revealed that these anchor residues are accommodated by a number of pockets in the floor of the binding groove (Garrett *et al.*, 1989; Rammensee *et al.*, 1993; Stern *et al.*, 1994).

1.1.8 Non-conventional processing pathways of exogenous antigens by MHC class I molecules

Although the MHC class II processing pathway is considered to be the most important pathway for the presentation of exogenous antigens, it is now apparent that exogenous antigens can also enter the conventional MHC class I processing pathway for presentation to $CD8^+$ T cells in a process called cross presentation (Ackerman and Cresswell 2004; Trombetta and Mellman 2005). There are a number of mechanisms that have been suggested to account for cross presentation including the 'leaking' of

antigens from endosomes into the cytosol and the delivery of antigens to the ER with subsequent escape into the cytoplasm. Most recently, it has been suggested that cross priming may occur as a result of ER incorporation into the phagosomal membrane during phagocytosis. Analysis of phagosomes revealed that several ER components including TAP and tapasin might be present in the phagosomal membrane (Garin *et al.*, 2001; Gagnon *et al.*, 2002). This discovery led to the proposal that during phagocytosis of bacterial pathogens, incorporation of ER membrane into the phagosome allows access of MHC class I molecules and other associated molecules such as the proteasome to peptide fragments within the phagosome (Watts 2002; Ackerman and Cresswell 2004).

In addition to conventional class I molecules, the non-classical MHC class I molecules belonging to the CD1 family are also thought to present peptides derived from exogenous antigens to CD8+ T cells (Braud *et al.* 1999; Kerksiek and Pamer 1999; Sugita *et al.*, 2004; Watts 2004). CD1 molecules are MHC-like molecules which have a non-polymorphic alpha chain encoded outside the MHC. Like classical MHC class I molecules, the alpha chain associates with β2M. CD1 molecules patrol the endocytic pathway for lipid antigens that they present to $CD8^+$ T cells. In humans, there are five CD1 loci (a-e), four of which are expressed, and recycle through different compartments of the endocytic compartment. The contribution of these different processing pathways towards immunity in many types of infection is unclear. However, studies have implicated CD1 molecules in protection against *M. tuberculosis* (Porcelli 1995; Stenger *et al.*, 1997; Stenger and Modlin 1998; Braud *et al.*, 1999; Rosat *et al.*, 1999).

1.2 Identification of naturally processed and presented $CD4^+$ T cell epitopes from exogenous antigens

Knowledge of the peptides selected for presentation to T cells from antigens and the underlying mechanisms of antigen processing and presentation is critical for the design of new 'epitope' based vaccines. Several approaches have been employed to identify or predict $CD8^+$ and $CD4^+$ T cell epitopes in antigens (Walden 1996; Lauemoller *et al.*, 2000; Schirle *et al.*, 2001). These include T cell mapping experiments that directly identify T cell epitopes and predictive methods that identify

peptides that may potentially be T cell epitopes. Predictive methods involve either the prediction of peptides that will bind to a given MHC allele based upon peptide motif information or the identification of naturally processed and presented peptides. With predictive methods, experimental verification is required to confirm that the peptide is a T cell epitope. This involves the analysis of MHC binding properties using synthetic peptides and T cell recognition using proliferation assays.

1.2.1 $CD4^+$ T cell mapping studies

One approach to identify epitopes within antigens is to examine T cell responses of either animals (e.g. mice) or individuals immunised with whole antigen (reviewed in Walden *et al.*, 1996). This approach usually employs the synthesis of overlapping peptides that cover the whole sequence of the protein. The peptides are then tested for their capacity to stimulate T cells that have been isolated from immunised donors. This technique of using T cells has the advantage that it is extremely sensitive, with $CD4^+$ T cells being able to detect even a single peptide-MHC complex on the surface of antigen presenting cells (Krogsgaard *et al.*, 2005).

Experiments using T cells isolated from immunised humans has led to the identification of a number of 'universal epitopes' in tetanus, diphtheria and mycobacterial antigens (Diethelm-Okita *et al.* 1997; Diethelm-Okita *et al.*, 2000; Al-Attiyah *et al.*, 2003; Caccamo *et al.*, 2004). Universal epitopes are regions of antigens to which the majority of humans in a population make a T cell response. For example, >80% of a panel of HLA-diverse individuals immunised with tetanus toxoid (TTD) had T cells that proliferated to peptide TTD $_{950\text{-}969}$ (Diethelm-Okita *et al.*, 2000). In addition, in-vitro T cell responses elicited by universal epitopes have been shown to be comparable to that of the whole antigen. The identification of these universal epitopes within an increasing number of antigens means that they could be included in synthetic epitope based vaccines to generate a protective immune response. Alternatively, universal epitopes could be used to boost the immune response to other immunising antigens. However, to date, there have only been a handful of antigens analysed for human T cell epitopes.

A major drawback in these studies is that they require a large number of individuals to be immunised with the antigen of interest. This is difficult because

even modified antigen can be potentially too toxic to risk immunisation of volunteers. In addition, murine epitopes provide limited information because they may not necessarily be the same as epitopes presented by their human counterparts due to differences in MHC molecules and antigen processing between the two species. In addition, T cell mapping studies provide no information of peptide flanking residues that are important in peptide binding and T cell activation (Moudgil *et al.* 1998).

1.2.2 Prediction of MHC binding

The identification of peptide binding motifs together with the development of MHC peptide binding algorithms, have permitted the prediction of both $CD8^+$ and $CD4^+$ T cell epitopes in a number of defined antigens (reviewed in Laumoller *et al.*, 2000 and Schirle *et al.*, 2001). In order to predict peptide binding motifs for MHC ligands, several groups have carried out MHC-peptide binding assays (Ruppert *et al.*, 1993; Hammer *et al.* 1994; Hammer *et al.* 1994; Kubo *et al.* 1994; Parker *et al.*, 1994). Other groups have identified sequence motifs based upon sequence information of natural ligands eluted from class II molecules (Falk *et al.*, 1991; Hunt *et al.*, 1992; Hunt *et al.*, 1992; Corr *et al.*, 1993; DiBrino *et al.*, 1993; Falk *et al.*, 1994).

Whereas the prediction of T cell epitopes has been successful for $CD8^+$ T cells, it has had only limited success in identifying epitopes for $CD4^+$ T cells. It is thought that one reason for this difficulty is the occurrence of class II peptides as nested sets, with peptide motifs not being confined to the core sequence like class I peptides (Rudensky *et al.*, 1991; Hunt *et al.*, 1992; Chicz *et al.*, 1993). Thus, bound peptides are not in register that complicates interpretation of pool sequencing experiments (Falk *et al.*, 1994; Rammensee 1995). Secondly, class II binding is often degenerate, with peptide binding motifs providing an incomplete description of the full diversity of peptides that are able to bind to MHC class II molecules (Panina-Bordignon *et al.*, 1989; Busch *et al.*, 1990; O'Sullivan *et al.*, 1990; Roche and Cresswell 1990; O'Sullivan *et al.*, 1991). It is known that peptides with very different sequences sometimes bind equally well to many different class II molecules. Analysis of the structure of DR / peptide complexes has led to the discovery that this

is because sequence independent interactions alone substantially stabilise class II / peptide binding (Stern *et al.*, 1994). In addition, positive interactions in some pockets can overcome negative interactions in others. As a consequence, all peptides have the potential to bind provided that their peptide side chains can be adequately accommodated in pockets in the class II groove. In addition, peptides predicted to bind to specific class II molecules may not be naturally processed and presented. As a result, it is difficult to predict T cell epitopes from antigens based upon sequence information alone.

1.2.3 Analysis of naturally processed and presented peptides

In the last 15 years, it has become possible to identify both $CD8^+$ and $CD4^+$T cell epitopes from antigens of interest by analysing naturally processed and presented peptides. This approach has the advantage over the other methods of T cell epitope identification in that it identifies naturally processed and presented peptides that represent the sum of all the selection events in the production of T cell epitopes. This is in contrast to predicting MHC class II binders that is based solely upon sequence information. However, there are also disadvantages in this technique in that it is a very laborious and time consuming approach and requires vast amounts of raw materials.

To date, the identification of naturally processed and presented bacterial derived T cell epitopes following infection has only been described for $CD8^+$ T cells. The first of these studies identified a naturally processed and presented *Listeria monocytogenes* derived murine $CD8^+$ T cell epitope bound by H2-K^d (Pamer *et al.*, 1991). In this study, mice where experimentally infected with *L. monocytogenes*, the splenic material isolated and the MHC class I peptides eluted for purification by HPLC and detection by CTL targeting activity. Subsequent studies identified 3 more naturally processed and presented $CD8^+$ T cell epitopes bound by H2-K^d, all of which were derived from secreted proteins important in *L. monocytogenes* virulence (Pamer 1994; Sijts *et al.*, 1996; Busch *et al.*, 1997). One of these was biochemically characterised by partial sequence analysis and database searching (Pamer 1994). A similar strategy was later used to identify $CD8^+$ T cell epitopes derived from *M. tuberculosis* (Flyer *et al.*, 2002). In this study, cell lines were experimentally

infected with *M. tuberculosis* and the MHC class I molecules isolated. Mass spectrometric sequencing of MHC class I isolated peptides led to the identification of three $CD8^+$ T cells from the same mycobacterial antigen.

The identification of bacterial derived $CD4^+$ T cells following experimental infection has proved difficult. One reason is because antigenic peptides of interest eluted from MHC class II molecules occur in very small amounts amongst the heterogeneous sea of peptides eluted from class II molecules (Sanderson *et al.*, 1995). This complex mixture makes the purification of antigenic peptides by HPLC a difficult task. In addition, the occurrence of class II peptides as nested sets complicates analysis. Fractionation of class II associated peptides leads to the dilution of individual class II peptides, with individual members of a nested set being eluted into different fractions.

A more common strategy in the identification of $CD4^+$ T cell epitopes has been to pulse large numbers of antigen presenting cells with huge amounts of antigen and subsequently analyse peptides eluted from the HLA molecule of interest using biochemical techniques. In theory, any antigen can be analysed in this manner and naturally processed and presented peptides bound by any MHC class II molecule studied. The first biochemical success was reported in 1992 when Nelson *et al.* identified several HEL-derived peptides amongst $I\text{-}A^k$ associated peptides. Since this study, there have only been seven more successes to date that are summarised in table 1.1 (Nelson *et al.*, 1992; Vignali *et al.*, 1993; Phelps *et al.*, 1996; Peakman *et al.*, 1999; Nepom *et al.*, 2001; Kaliyaperumal *et al.*, 2002; Dengjel *et al.*, 2004; Meiring *et al.*, 2005). The majority of these studies have analysed peptides derived from antigens implicated in the development of auto-immune disease bound to human DR molecules expressed by lymphoblastoid B cell lines. Typically, these studies have used in the order of 10^{10} cells and milligrams of antigen for success. However, the most recent report analysed peptides derived from the meningococcal outer membrane protein Por A presented by primary human dendritic cells. In this study, only 10^6 antigen presenting cells were required for successful identification of several T cell epitopes. This reduction in the number of antigen presenting cells is probably due to advancements in biochemical techniques over the past decade.

Biochemical advancements will be discussed together with the problems still faced in these studies in the following section.

Year	Authors	Antigen / MHC molecule/ APC	Method of epitope identification and characterisation
1992	Nelson and Unanue	HEL / IA^k / murine B cells	• Identification by subtractive methods using HPLC tracings • Characterisation by Edman pool sequencing of fractions containing antigen derived peptides and mass matching to HEL sequence
1993	Vignali *et al.*	HEL / IA^k / murine B cells	• Identification by testing the ability of individual HPLC fractions to stimulate T cell hybridomas • Characterisation by MALDI-MS analysis of stimulatory fractions and mass matching
1996	Phelps *et al.*	α3 (IV) / DR15 / EBV B cells	• Identification by subtractive methods using MADLI-MS • Characterisation by mass matching to α3 (IV)/ sequence
1999	Peakman *et al.*	IA-2 / DR4 / EBV B cells	• Identification by subtractive methods using MALDI-MS • Characterisation by MS/MS analysis of 'extra' antigen derived peptides and mass matching
2001	Nepom *et al.*	GAD65 expressing DR4 / EBV B cells	• Identification and characterisation by online nano-flow RP-LC-MS followed by MS/MS analysis of all peptides and mass matching
2002	Kaliyaperumal *et al.*	Chromatin / A20 / EBV B cells	• Identification by stimulation of T helper clones with HPLC fractions • Characterisation by MS/MS analysis of active fractions and mass matching
2004	Dengjel *et al.*	Cyclin D1 / HLA-DR4 / EBV B cells	• Identification by subtractive methods using MALDI-TOF MS • Characterisation by MS/MS analysis and mass matching
2005	Meiring *et al.*	Meningococcal porin A / DR1 / human dendritic cells	• Identification by online nano-flow RP-LC-MS to identify ^{15}N labelled peptides • Characterisation by MS/MS analysis and mass matching

Table 1.1 Reported biochemical successes in the identification of T helper cell epitopes in exogenous antigens

The table shows the 8 published accounts of identification and characterisation of $CD4^+$ T cell epitopes from exogenous antigens. Methods of characterisation include Edman sequencing, mass matching and most recently, mass spectrometric sequencing (MS/MS) using tandem mass spectrometry. In order to reduce the number of peptides to be characterised, groups have either used subtractive methods or the testing of HPLC fractions with T cell clones / hybridomas for their stimulatory capacity.

1.2.3.1 Problems currently faced in biochemical studies

Problems in biochemical studies have arisen due to the complexity of peptides eluted from class II molecules and due to limitations in the sensitivity of sequencing analysis. Vignali *et al.*, (1993) estimated that individual antigen derived peptides represent less than 1% of the total peptide pool isolated from MHC class II molecules which typically consists of thousands of peptides. As a consequence, the main problems have been (i) identifying peptides of interest within the complex peptide pool eluted from MHC class II molecules and (ii) being able to characterise extremely small amounts of these peptides. In the first published success, Nelson *et al.*, (1992) used Edman degradation to sequence peptides that at the time required approximately 10 pmoles of pure peptides for success. This study utilised 80g of HEL in 80 litres of supernatant containing 10^{10} cells and was successful due to the dominance of certain antigen derived peptides that permitted their sequencing despite partial purification. However, in subsequent studies carried out by Vignali *et al.*, (1993) and Phelps *et al.*, (1996), peptides were present at very small amounts and were therefore unable to be sequenced. In both these cases, peptides were characterised by mass matching.

However, recent advancements in mass spectrometry have largely overcome the problem of sequencing, with modern day mass spectrometers being able to detect sub-femtomole amounts of sample (Lemmel and Stevanovic 2003). These tandem mass spectrometers provide MS/MS data that permits the partial sequence analysis of peptides of interest. Peakman *et al.*, (1999) was the first group to use MS/MS

sequencing to characterise antigen derived peptides, with sensitivity of analysis being estimated to be at the femtomolar level. Since this study, there have been further advancements in sensitivity, with Nepom *et al.*, (2001) sequencing a peptide estimated to occupy only 10 MHC class II molecules per cell on the 10^7 cells analysed (~1.7 attomoles) by nanospray-Fourier transform mass spectrometry. These advances in mass spectrometric sensitivity have substantially eroded the former principal hurdle in MHC peptide analysis of peptide characterisation. As a consequence, identifying peptides of interest within the complex peptide pool for sequence analysis remains the major problem.

One solution has been to identify T cell epitopes within the complex peptide pool without fractionation (Nepom *et al.*, 2001; Meiring *et al.*, 2005). This has been made possible by the recent development of highly sensitive mass spectrometers. However, this method is time consuming and it is very expensive to generate sequence information for all peptides eluted from class II molecules. Alternative strategies have been to reduce the complexity of class II peptides and thus simplify the identification of potential T cell epitopes. In the first approach, T cells have been used to identify fractions in which epitopes are present by testing the stimulatory capacity of each fraction (Vignali *et al.*, 1993; Kaliyaperumal *et al.*, 2002). In these studies, the use of T cells reduced the problem to identifying a few peptides in a hundred rather than a few peptides in thousands. In the second approach, subtractive methods are employed in which comparisons of MHC class II peptides eluted from both sham pulsed and antigen pulsed antigen presenting cells are made to identify extra 'putative' peptides (Nelson *et al.*, 1992; Phelps *et al.*, 1996; Peakman *et al.*, 1999; Dengjel *et al.*, 2004). Nelson *et al.*, (1992) identified major HEL derived epitopes by comparing the UV tracings from both antigen and sham pulsed cells. This was made possible by the abundance of HEL derived peptides. However, in all other reports, only minor differences have been observed in the overall UV profiles obtained from antigen pulsed and control cells. This complexity is illustrated in figure 1.5 which shows UV tracings obtained from the study carried out by Phelps *et al.*, (1996) in the identification of naturally processed and presented peptides derived from the type IV collagen chain in the α3 antigen (α3 (IV)). As a result of the complexity of class II peptides following fractionation, MALDI-TOF MS was

employed in this study and others to identify antigen derived peptides. Mass spectrometric analysis has shown that up to 100 peptides can typically elute in one HPLC fraction, making analysis very time consuming and difficult when trying to identify the scarce antigen derived peptides by subtractive methods.

Clearly, it would be of major advantage in both of these methods to be able to unambiguously identify antigen-derived peptides in complex peptide mixtures. Conventional protein labelling techniques used in antigen processing studies have included radioactive labelling as well as discrete labels such as biotin and colloidal gold. However, these labelling techniques have limited utility when applied to tracing proteins through environments in which protein degradation occurs, such as through the antigen processing system of antigen presenting cells. This is because discrete labels are introduced at a limited density, so may not be contained within important fragments. Moreover, the labelling process usually modifies specific amino acids with potential effects upon biological activity, such as binding to class II molecules. Metabolic labelling, usually with radioactive isotopes, overcomes the deficiencies of discrete labels. However, it has the disadvantages of being potentially hazardous, and difficult to control in large scale experiments required to generate sufficient peptide for biochemical analysis.

The use of stable isotopes in protein labelling overcomes the deficiencies of both radioactive and discrete labels and will be discussed in the following section.

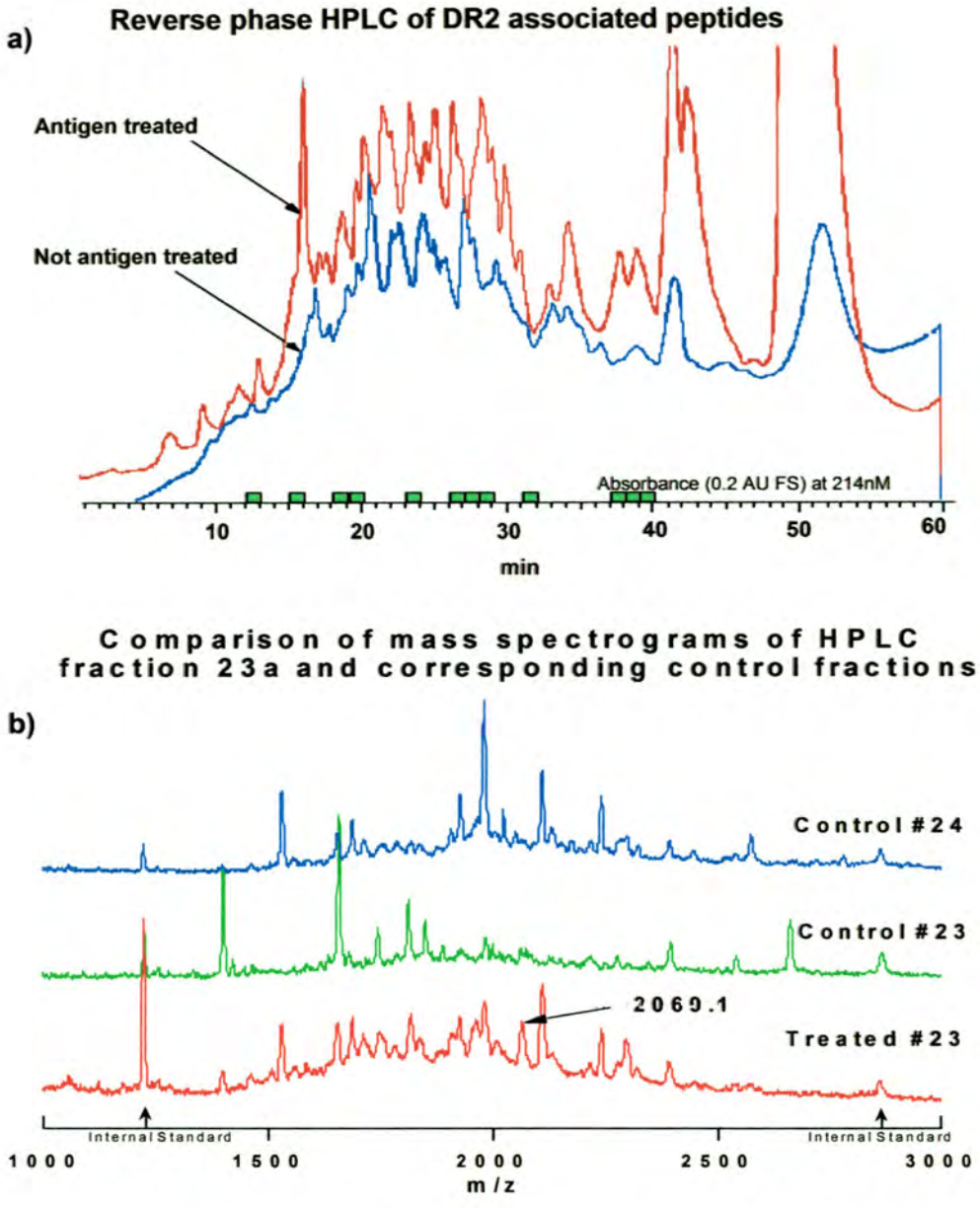


Figure 1.5 Identification of naturally processed and presented peptides from α3 (IV)

(a) Shows a UV spectrogram of a RP HPLC separation of MHC class II peptides from α3 (IV) pulsed and sham pulsed cells. The tracing is extremely complex with only minor differences being observable. **(b)** Mass spectrograms of Fr 23 from the HPLC separation of antigen pulsed cells and Fr 23 and 24 from the HPLC separation of sham pulsed cells. An extra antigen derived peptide (MW 2069.1) is shown in fraction 23 of the HPLC separation of the eluted antigen pulsed peptide pool which is absent from the control fractions.

1.2.4 Advances in biochemical techniques

The final section of the introduction will briefly review the biochemical techniques upon which the analysis of MHC class II eluted peptides is dependent, namely mass spectrometry and HPLC. The most recent advancements in these biochemical techniques will be reviewed and the potential use of these 'new' techniques in simplifying biochemical analysis of complex peptide mixtures highlighted. These advancements include in mass spectrometric-based protein labelling and the development of multidimensional liquid chromatography. Such advancements have great suitability for antigen processing studies in which thousands of peptides need to be efficiently separated prior to MS analysis.

1.2.4.1 Mass spectrometry

Mass spectrometry (MS) is a technique used to accurately measure the mass to charge ratio (m/z) of ions. It is routinely used to identify unknown proteins either by the mass of the protein itself or by the masses of its fragments, to determine the structure of compounds and to quantify compounds in samples of interest (Hamdan and Righetti, 2002; Aebersold and Mann, 2003; Sommerer *et al.*, 2006). The development of two 'soft' ionisation techniques in the 1980's known as electrospray ionisation (ESI) and matrix-assisted laser desorption ionisation (MALDI) led to the application of mass spectrometry to many biological problems including the identification of proteins (Siuzdak 1994). This is because 'soft' ionisation techniques allow large analytes, for example proteins, to become ions without significant fragmentation, thereby making their molecular weight obtainable using mass spectrometry. Soft ionisation techniques also give rise predominantly to a singly positively charged parent ion (M+1) during mass spectrometric analysis, making them ideal for many applications including the identification of large macromolecules.

In MALDI-MS, a laser is used to ionise the sample, generating positively charged gaseous ions. The sample is mixed with a substance called the 'matrix' that absorbs the laser energy and transfers it to the sample, thereby preventing fragmentation of the analyte. In ESI, the analyte is passed through a capillary held at

high potential, leading to evaporation of the droplets and the production of charged gaseous ions. Following ionisation of the analyte by one of these soft ionisation techniques, the ions are separated in a mass analyser according to their mass / charge ratio (m/z) and the then recorded by a detection system. This information is acquired under the control of a data processing system that generates a mass spectrogram, from which the mass of the ion is obtained. The detector records the signal as an electron cascade, which is displayed on an oscilloscope, and then the data is shown graphically as m/z (x axis) against the number of ion hits (y axis). Since charge (z) is equal to +1, then m/z is equivalent to the mass of the ion. Commercially, both MALDI and ESI ion sources are found in conjunction with different types of mass analysers. These mass analysers include time of flight (TOF), ion trap (IT), quadropole, and fourier transform ion cyclotron resonance (FT-ICR).

1.2.4.1.1 Matrix-Assisted Laser Desorption Ionisation – Time Of Flight (MALDI-TOF) mass spectrometry

Figure 6 shows a diagram of a MALDI-TOF mass spectrometer and illustrates how such an instrument works. In MALDI-TOF, the ions are accelerated towards the detector along a flight vessel at a speed determined by the mass of the ion, with lighter ions reaching the detector first. The mass of the ion is calculated from the drift time and is shown in figure 1.7. Since almost all ions possess only a single positive charge, time of flight is proportional to the square root of the mass of the ion.

Time of flight mass spectrometry works on the principle that if ions are accelerated with the same potential from a fixed point at a fixed initial time and are allowed to drift, the ions will separate according to their mass to charge ratios. However, when accelerated, ions exhibit a broad spread of energy, causing them to travel towards the detector at different speeds. Thus, ions of the same mass reach the detector at slightly different points in time. The use of Delayed Extraction technology overcomes this initial distribution of energies, with ions being formed in a weak electrical field, and then a high voltage pulse applied after a predetermined time decay to accelerate the ions. Delayed extraction is used to improve mass accuracy.

Once the ions have reached the end of the flight tube, they are detected in either linear or reflector mode. In the latter, an electrostatic mirror is used to reflect the ions during their course, causing them to travel a further distance prior to detection. This longer drift time gives rise to greater mass accuracy during detection. Development of the 'reflector' mode has enabled individual isotopic peaks of peptides to be visualised during mass spectrometric analysis. Prior to this, only one peak for a peptide could be identified during linear detection. The visualisation of individual isotopic peaks during mass spectrometric analysis has paved the way for the isotopic labelling of proteins that is currently used to identify and quantify compounds in complex samples.

1.2.4.1.2 Tandem mass spectrometry

Soft ionisation techniques do not lead to the fragmentation of proteins or peptides. As a consequence, only mass information is acquired during analysis using soft ionisation techniques. In order to provide structural information, many mass spectrometers available upon the market are now tandem mass spectrometers in which peptides or proteins are fragmented to obtain sequence or structural information (Fenselau *et al.*, 1993; Lemmel and Stevanovic, 2003). These instruments include triple quadrupole, quadrupole ion trap, quadrupole time of flight instruments and MALDI-(TOF/TOF) instruments.

In tandem mass spectrometry, a first mass analyser is used for selection of the desirable ion. The ion is then subjected to fragmentation by collision with an inert gas inside a 'collision chamber'. This is called low-energy collision induced dissociation (CID). A second mass analyser is then employed to separate out fragment ions based upon their m/z values. From these fragments, a partial sequence can be determined for peptides using known information of the expected fragmentation patterns. This technique is extremely powerful and is commonly used in conjunction with HPLC to identify unknown proteins in extremely complex mixtures.

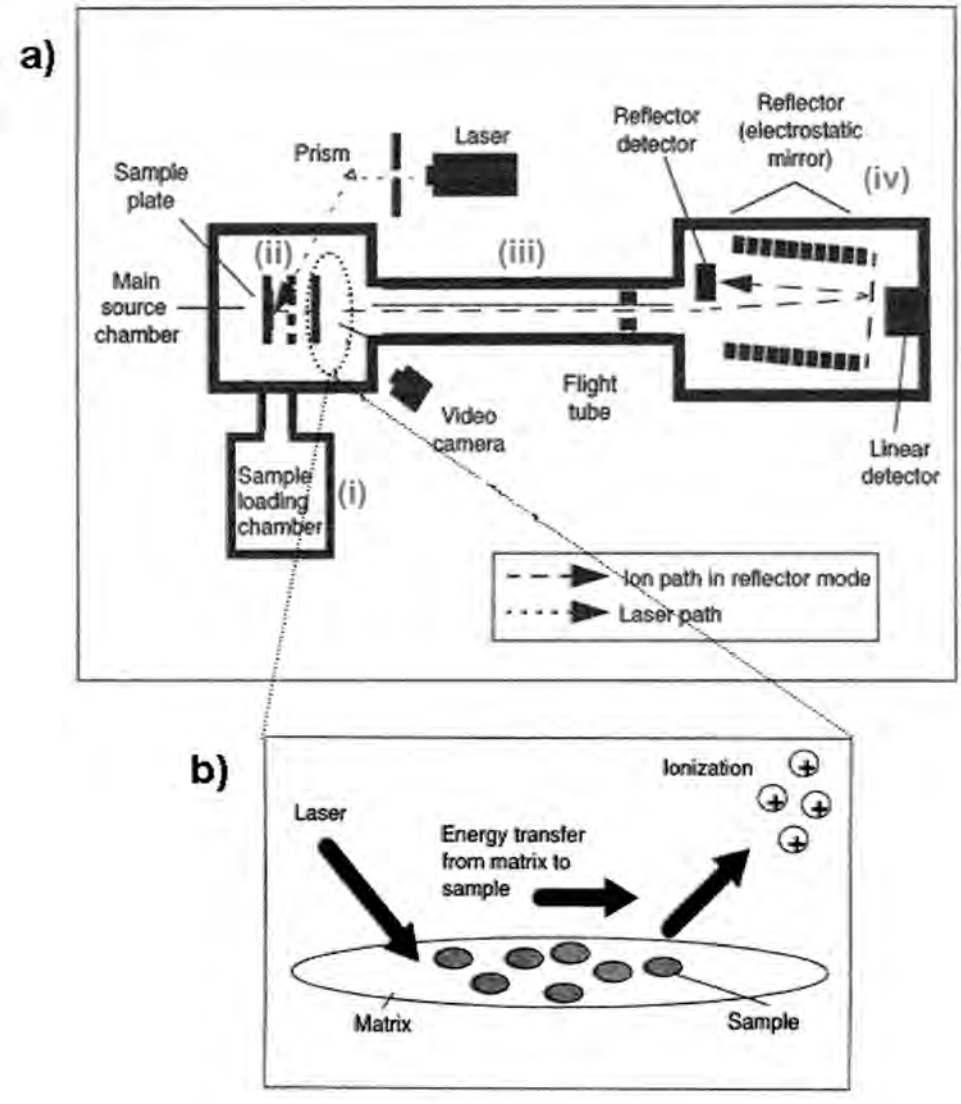


Figure 1.6 Diagram of a MALDI-TOF mass spectrometer

Figure adapted from Voyager ™ Biospectrometry ™ Workstation User's Guide

(a) Shows how a MALDI-TOF mass spectrometer works

(i) The sample plate is placed into a loading chamber.

(ii) Upon loading, the sample plate is taken to a main source chamber where the sample is ionised by a laser.

(iii) Positive ions are subsequently propelled along a flight tube to a detector

(iv) Detection can either be in linear mode or reflector mode. In the latter, the path of the ions is deflected using an electrostatic mirror, increasing their flight distance prior to detection.

(b) The sample is mixed with a component called the matrix that absorbs the laser energy and transfers it to the analyte molecules, leading to ionisation without fragmentation of the analyte.

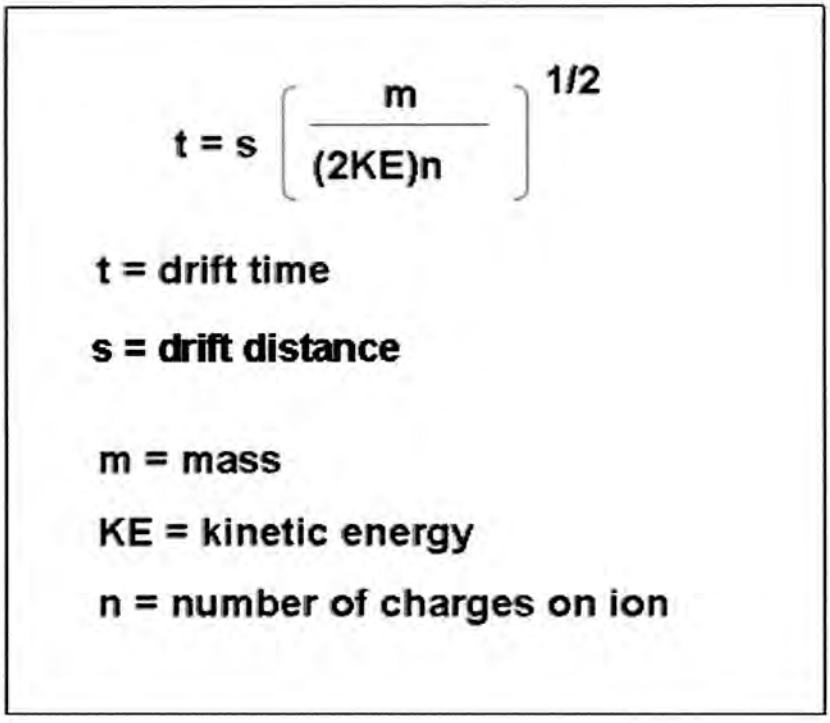


Figure 1.7 Equation used to derive ion mass

(Taken from Voyager ™ Biospectrometry ™ Workstation Users's Guide)

The time taken for the ions to reach the detector at the opposite end of the flight tube (drift time) is proportional to the square root of the mass of the ion and is calculated using the equation above. Approximate ion mass is determined in this way.

1.2.4.2 Mass accuracy and calibration

Unless a tandem mass spectrometer is employed, structural information is not obtained for an unknown peptide / protein species. It is therefore paramount that accurate masses are obtained during the acquisition of data. The mass of the ions is measured in Daltons in mass spectrometry. Mass accuracy indicates the accuracy of the mass information provided by a mass spectrometer. It is the difference which is observed between the theoretical mass and the measured mass and is expressed in parts per million (ppm). Equations explaining the derivation of mass accuracy are shown in figure 1.8a. Mass accuracy is linked to the resolution capability of the mass spectrometer in use and represents the ability to separate two adjacent masses. Thus, the mass of an ion can only be measured as accurately as the instrumentation allows.

For optimum mass accuracy, the mass spectrometer is calibrated using standards of known mass. The mass spectrometer should be calibrated at least once a day but most commonly, calibrants are included during the analysis of each sample. A calibration file is generated based upon the observed time of flight of masses of known calibration standards. The file is generated using standard mixtures containing components with masses throughout the range of interest. The calibration file is then used to adjust the masses of unknown peptide / protein species in the sample of interest, thus correcting for any mass inaccuracies resulting during the acquisition of data. The calibration software compensates for changes in accelerating voltage during acquisition that may lead to ions of the same mass / charge ratio reaching the detector at slightly different times, thus compromising mass accuracy. The general calibration equation employed by the mass spectrometer used is this project in shown in fig 1.8b.

The calibration function allows the generation of two types of calibration: external and internal calibration. An external calibration file is generated when the standards are not mixed with the sample and are therefore analysed separately. In contrast, an internal calibration file is generated when the standards are mixed with the sample and thus analysed at the same time. Internal calibration provides better mass accuracy compared to external calibration and is commonly used for the calibration of unknown peptide masses in the identification of proteins in MS peptide finger printing (Sommerer *et al.*, 2006).

(a)

Mass accuracy (Δm) = mass real – mass measured

It is often expressed in parts per million (ppm):

ppm = 10^6 * Δm accuracy / mass measured

i.e.: theoretical mass: 1000, measured mass: 999.9 error: 100 ppm

$$t = t_o + A\sqrt{m/z} + (\text{higher order terms})$$

Where:

t = time of flight

t_o = Difference in time between the start time of the analysis and the time of ion extraction in Delayed Extraction,

$$A = \frac{\text{Effective Length (mm)}}{\sqrt{\text{Accelerating Voltage (Kv)}}} \times \sqrt{\frac{m_o}{e}} \times 10^9$$

Where:

m_o = 1 dalton mass in SI units (1.66054×10^{-27} kg)

e = charge of electron in SI units (1.602177×10^{-19} coulomb)

Effective length = length of flight tube corrected for an ion acceleration through the stages of the source and flight tube, and for the impact of the guide wire.

m/z = mass-to-charge ratio

Figure 1.8 Mass accuracy and calibration equations

(Taken from Voyager ™ Biospectrometry ™ Workstation Users's Guide)

The above figure shows the derivation of the equations for mass accuracy and calibration. **(a)** Mass accuracy is the difference between the theoretical mass and the measured mass and it is expressed in parts per million (ppm). **(b)** To obtain accurate ion masses, a calibration procedure using reference standards of known mass is used. For each standard, values for t_0 and A are calculated which are based upon the configuration of the system such as flight tube length and accelerating voltage. The calibration file created from the time of flight of known standards is then applied to mass spectra containing ions of unknown mass. This calibration software compensates for changes in accelerating voltage that is used to propel the ions along the flight tube.

1.2.4.3 Use of stable isotopes in protein labelling

Stable isotopes (e.g. ^{15}N, ^{18}O and ^{13}C) are being increasingly used in proteomics for protein quantification and identification (Chen *et al.* 2000; Hamdan and Righetti 2002; Aebersold and Mann 2003; Tao and Aebersold 2003). Stable isotopes are introduced into proteins by either metabolic labelling with heavy salts / aminoacids / carbohydrates; or enzymatically via the transfer of ^{18}O from water to peptides or by chemical reactions. This technique is based upon the fact that chemically identical analytes of different stable-isotope composition can be differentiated in a mass spectrometer based upon either a difference in mass (^{15}N, ^{18}O) or overall isotopic pattern (^{13}C). The metabolic labelling of antigens with stable isotopes overcomes the short comings of both discrete labels and radioactive labelling as they are both safe to use and are incorporated into the framework of the antigen, preventing the labelling process from interfering with processing and presentation of the antigen.

To date, however, stable isotopes have been used in few studies analysing naturally processed and presented peptides (Lemmel *et al.,* 2004; Ringrose *et al.*, 2004; Meiring *et al.,* 2005). The only example of the use of stable isotopes in the identification of $CD4^+$ T cell epitopes was provided by Meiring *et al.,* (2005) who identified a number of $CD4^+$ T cell epitopes from the meningococcal outer membrane protein porin A (Por A) by analysing MHC class II eluted peptides. In this study, bacteria were cultured in minimal media containing either $^{14}NH_4Cl$ or

$^{15}NH_4Cl$, and then the outer membrane vesicles (OMV) containing Por A subsequently isolated and used to pulse dendritic cells in a 1:1 ratio. The MHC class II peptides were subsequently isolated, separated by HPLC and labelled peptides identified as 'spectral couplets' during mass spectrometric analysis. These couplets are generated by labelled peptides increasing in mass following the incorporation of ^{15}N.

1.2.4.3.1 Carbon-13 labelling technique

Whereas 100% labelling with ^{15}N leads to a discrete change in mass, the overall isotopic pattern of peptides can be changed by labelling with 2.3% ^{13}C (Zou *et al.*, 2004). Therefore, instead of distinguishing spectral couplets in complex spectra, peptides labelled with carbon-13 should be distinguishable from all other peptides in the spectra. This approach is based upon the fact that there are two stable isotopes of carbon, carbon-12 (^{12}C) and carbon-13 (^{13}C). Carbon-12 represents 98.892% of all naturally occurring carbon atoms whilst ^{13}C represents 1.08%, giving rise to peptides with characteristic isotopic patterns.

For a given peptide, the first peak represents detection of ions of peptides with zero carbon 13 atoms, the second represents the number of molecules with one ^{13}C atom, and the 3rd represents those with two ^{13}C atoms and so forth. The relative abundance of these peptide ions depends on the proportion of ^{13}C and the number of C atoms within the peptide, and can be calculated using the binomial distribution (figure 1.9). The analysis suggests that ^{13}C enriched peptides should be clearly distinguishable from peptides containing the usual abundance of ^{13}C by differences in their isotope pattern, provided account is taken of their length and the likely range of carbon atoms they contain. Thus, this labelling technique has potential to produce labelled peptides that are easier to identify because they appear visually distinct to unlabelled peptides. Although this approach has not been used in antigen processing studies, experimental work with the test protein Schistomal glutathione transferase (sGT) has shown this ^{13}C labelling technique to be a highly sensitive and reproducible method of identification of labelled peptides in complex mixtures.

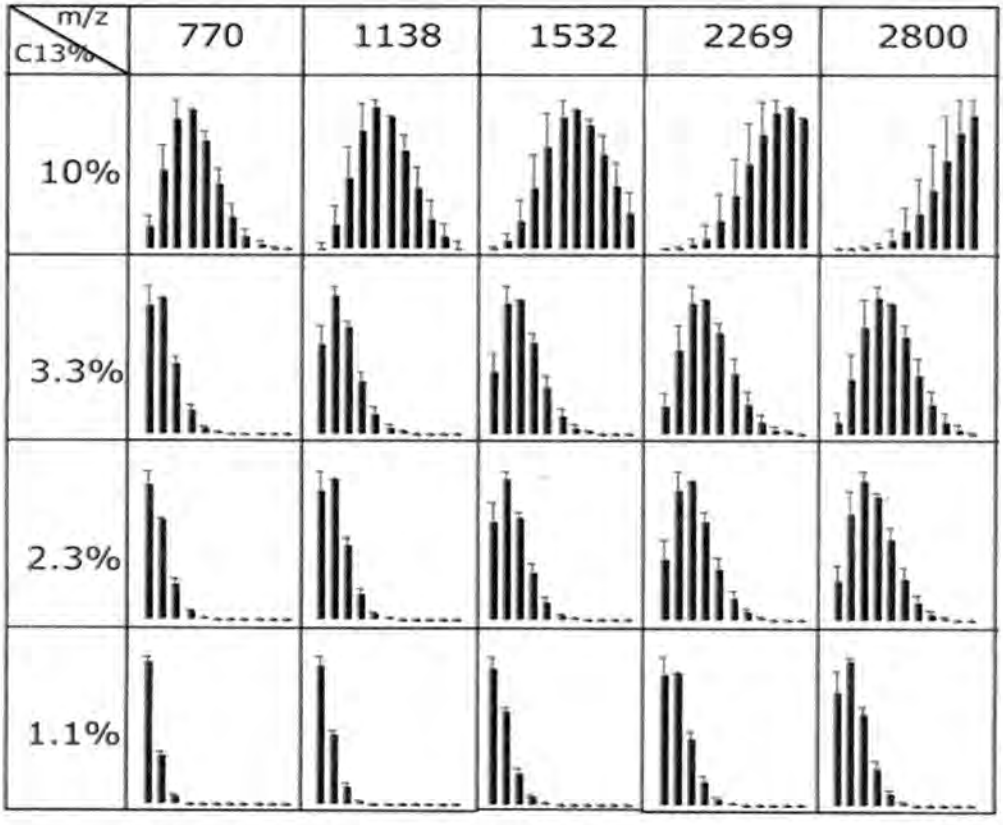


Figure 1.9 Predicted isotopic patterns for peptides of different mass at four different carbon-13 abundances

Each peptide species in a mass spectrogram is made up of several isotopic peaks, one mass unit apart, due to peptide ions containing 0, 1, 2 etc. atoms of carbon-13. The expected relative abundance of peptide ions containing 0-4 carbon-13 atoms is calculated from the binomial distribution. The number of carbon atoms and the abundance of carbon-13 influence the pattern. Error bars indicate the degree of variation in the patterns seen with peptides containing 1.96 standard deviations more or fewer carbon atoms compared to the mean number of carbon atoms per dalton found in human proteins. Vertical comparison demonstrates that carbon 13-enhanced peptides have distinct isotope patterns at each of the representative masses (Zou *et al.*, 2004).

1.2.4.3 High Performance Liquid Chromatography (HPLC)

HPLC is commonly employed to separate, purify, identify and quantify chemical substances (Issaq *et al.*, 2002). Prior to the development of HPLC in the mid-seventies, most chemical separations were carried out using a variety of techniques including open-column chromatography, paper chromatography, and thin layer chromatography. However, these techniques were inadequate for the quantification of compounds and resolution between similar compounds. The use of high pressure or high performance liquid chromatography (HPLC) as it is commonly referred to has resulted in a reduction in flow through time and the purification time of compounds.

HPLC separation of complex peptide mixtures prior to mass spectrometric analysis is the common method of biochemical analysis in both proteomic studies and in the characterisation of MHC associated peptides (Lemmel and Stevanovic, 2003). In 1992, Hunt *et al.* first pioneered the use of liquid chromatography – MS/MS for the analysis of complex peptide mixtures by studying MHC class I associated peptides. Since this initial study, single dimensional chromatography has routinely been used prior to either tandem mass spectrometry or MALDI mass spectrometry in studies characterising both class I and class II peptides. It has, however, become apparent that one dimensional chromatography does not provide sufficient peak capacity to efficiently separate peptide mixtures as complex as those eluted from MHC molecules, or total cell lysates in the case of proteomic studies.

With the advent of proteomics, various combinations of peptide separation techniques have been explored in order to provide better separation of complex peptide mixtures (Wang and Hanash, 2003). The most popular of these is strong cation exchange followed by reverse phase chromatography. Whereas a 1D chromatographic system equipped with a standard reverse phase column can only efficiently resolve about one hundred peptides, a 2D separation system has the potential to resolve thousands of peptides (Michrom Biosciences, application note No. 111). As a direct result of the enormous improvement in the separation capability of these systems, multi-dimensional liquid chromatography has been

successfully employed in many proteomic studies, identifying thousands of proteins in proteomes isolated from both cells and pathogens (Nilsson, 2002; Wilson *et al.*, 2004). Although 2D HPLC has not yet been used in antigen processing studies, it has great potential to significantly improve the separation of peptides compared to traditional 1D separation techniques.

1.3 Purpose of project

The purpose of this project was to test whether the carbon-13 labelling technique could be used to identify peptides of immunological interest within complex mixtures. The ultimate aim of the work would be to use the carbon-13 labelling technique to identify naturally processed and presented MHC class II bound peptides derived from whole bacteria.

With this in mind, development of the approach initially employed the use of outer membrane proteins isolated from 3 different strains of *E. coli*. Using these extracts, work was carried to identify immunologically relevant outer membrane proteins and to produce carbon-13 labelled proteins. Following preliminary work, the project switched to using the single antigen tetanus toxin C fragment (TTCF) to simplify characterisation of carbon-13 labelled peptides following their identification in complex peptide mixtures. It was hypothesised that following validation of the technique using a single antigen, the more complex bacterial outer membranes could be subsequently analysed using the same technique. Using this antigen, work was carried out to identify fragments generated following the in-vitro digestion of TTCF using lysosomal extracts isolated from lymphoblastoid human B cells. This was carried out with the intention of subsequently going on to study naturally processed and presented TTCF peptides bound to MHC class II molecules at the cell surface.

2. Methods and materials

2.1 Cells and reagents

2.1.1 Bacterial strains and toxins

The bacterial strains used in this project were *E. coli* O18, K12 and O12. These strains were chosen from a bank of freeze-dried bacterial strains kindly provided by Professor Ian R Poxton's laboratory at Edinburgh University. An N-terminal histidine tagged tetanus toxin C fragment fusion was kindly provided by Professor Colin Watts at Dundee University. This fusion protein was encoded by pET 16B and expressed in the cell line *E. coli* strain BL21 (DE3).

2.1.2 B cell lines and cell culture conditions

The B cell line used in this study was the Epstein Barr virus (EBV) transformed human Pala lymphoblastoid B cell line that was very kindly provided by Professor Colin Watts at Dundee University. B cell media was made up using 1L of GIBCO RPMI 1640 medium supplemented with 5% fetal calf serum, 200 mM L-glutamine, 10,000 units / ml penicillin, and 10mg / ml streptomycin solution. The B cells were grown in 25ml of this medium and incubated at 37 °C with 5% CO2. The B cells were split 1 in 4 using fresh B cell media culture 3 times a week.

2.2 Methods – part 1

2.2.1 Carbon-13 labelling of bacterial outer membrane proteins

To produce carbon-13 labelled bacterial outer membrane proteins, the appropriate *E. coli* strain was grown in M9 minimal medium supplemented with carbon-13 glucose, as described previously (Zou *et al.*, 2004). The bacteria were initially grown in a starter culture of between 5-15 ml of Luria broth (LB) (10g tryptone, 5g yeast extract, 10g NaCl; per litre) and then diluted into either a larger volume of LB broth or into M9 minimal medium (see table 2.1 for recipe) to give a carry over of 3.3% of the initial starter culture. This amount of carry over of rich broth into minimal medium is negligible. The amounts of carbon-13 glucose used for

enrichment were 2.3%, and 3.3%, given as a percentage of the total amount of glucose in the growth medium. To obtain these percentages of carbon-13 in the growth media, 70mg (for 2.3%) and 100mg (for 3.3%) of carbon-13 labelled glucose was added.

M9 minimal medium (1L)
• 800ml H20 • 200ml 5*M9 (0.24M Na_2HPO_4, 0.11M KH_2PO_4, 0.043 M NaCl) • 1ml salt mix (4mM $ZnSO_4$, 1mM $MnSO_4$, 0.7mM H_3BO_3, 0.7mM $CuS0_4$) • 2ml 1M $MgSO_4$ • 2ml 50Mm $CaCl_2$ • 400µl 5mM $FeCl_3$ • 800µl thiamine (0.5mg/ml) • 9.4ml $NH4Cl_2$ • 15ml Glucose [20% (w/v)] • 1ml carbenicillin (50mg/ml) • 4.6ml ^{13}C-glucose (15mg/ml)

Table 2.1 Recipe for M9 minimal medium
% glucose calculated as (w/v)

2.2.2 Extraction of *E. coli* outer membrane extracts

Outer membranes were extracted from *E. coli* using the 'Sarkosyl' method (Filip *et al.*, 1973). Bacteria were initially grown to late log phase (OD_{550} ~0.8) in either LB broth or M9 minimal medium spiked with carbon-13 glucose and harvested by centrifugation at 10,000g for 15 minutes at 4°C. The cell pellets were subsequently washed 3 times in 50ml of 0.05M sodium phosphate buffer (0.05M $NaH_2PO_4.2H_20$ and 0.05M Na_2HPO_4; pH 7.4) containing 0.15M sodium chloride. Following each wash, the cell pellet was recovered by centrifugation at 10,000g for 15 minutes, carried out at 4°C using a Sorvall RC5B centrifuge. After the 3 wash

steps, the cells were resuspended in 9ml of pyrogen free water and disrupted by sonication which was carried out on ice for 1 minute, using an amplitude of 5-10μm (MSE Soniprep 150). The sonication step was carried out a total of 6 times, with the cells being left on ice for 30 seconds between each sonication step. To solubilise the outer membrane, 7% Sarkosyl solution was subsequently added to the cells to give a final concentration of 0.07%. The unbroken cells were then removed by centrifugation at 10,000g for 15 minutes, carried out at 4 °C using an Eppendorf 5810R centrifuge. The outer membrane proteins were subsequently recovered by centrifugation at 50,000g for 1 hour at 4°C, washed in 10 ml of pyrogen free water, and once again recovered by centrifugation at 50,000g for 1 hour at 4°C using a Sorvall OTD centrifuge. Following the wash step, the insoluble outer membrane pellet was re-suspended in 1ml of water and stored at -20 ° C until use.

2.2.3 Confirmation of carbon-13 labelling of bacterial proteins in outer membrane extracts

To confirm that the proteins were labelled and to determine the correct amount of labelling, the outer membrane extracts were run out on a 10% SDS-PAGE gel. Prior to SDS-PAGE analysis, the outer membrane proteins were boiled in 2% SDS for 3 minutes. Protein bands were then excised from the gel, sliced into 1mm cubes, and washed 3 times in water. They were then dried under vacuum in a speed vac (Savant AES 1000 / AES 1010). The dried gel pieces were subsequently reduced by incubating the gel pieces with 10mM DTT in 50-100μl of 100mM $NH_4(HCO_3)_2$ / 5% ACN for 45 minutes at 56°C. The gel pieces were then alkylated by incubating the gel pieces with 50mM iodoacetamide in 50-100μl 100mM $NH_4(HCO_3)_2$ in the dark for 2 hours. Following alkylation, the gel pieces were placed in a speed vac and dried and then 20μg of modified porcine trypsin (Promega) added in a 50mM tris-phosphate buffer (pH 7.4) to digest the protein for 4 hours at 37°C. After digestion, the peptides were extracted from the gel pieces by initial incubation with 100μl of 50% ACN / 50% $NH_4(HCO_3)_2$ for 10 minutes (repeated once) followed by incubation for a further 10 minutes with 100μl 5% formic acid / 50% ACN (repeated once). All of these extracts were pooled together and dried under vacuum in a speed vac to a volume of 10μl.

2.2.4 MALDI-MS sample preparation & data acquisition

To prepare the extracts for matrix assisted laser desorption ionisation – time of flight mass spectrometry (MALDI-TOF MS), 30µl of 0.1% trifluoroacetic acid (TFA) was added and the extracts briefly vortexed, and then sonicated for 15 minutes. The samples were concentrated using a Zip Tip™ (Millipore) and then eluted using saturated matrix solution (α-cyano-4-hydroxycinnamic / 70% ACN / 0.1% TFA). The eluted peptides were spotted directly onto a 400 well MALDI plate. The matrix, α-cyano-4-hydroxycinnamic (Sigma), co-crystalised with the peptide extracts by the dried droplet technique. The dried sample was then analysed in a Voyager DE Pro mass spectrometer in reflector mode with an extraction delay of 180ns, accelerating voltage of 20,000V, grid of 76%, and by examining the mass range 600-4000 Daltons. Per spectrum, 400-800 laser shots were collected, with all spectra being subsequently processed to remove electronic noise and to smooth the baseline. Peptides were calibrated using the two autolysis products of trypsin (842.51 Da and 2211.1 Da).

2.2.5 Determination of the amount of protein in outer membrane extracts

The amount of protein (mg/ml) in the outer membrane extracts was determined using the Lowry method (Lowry *et al.*, 1951). Between 10-50µl of the outer membrane extracts were added to 350µl of distilled water and then mixed with 1.2ml of 12.5% $NaCO_3$ and 0.2ml of 0.1% $CuSO_4.5H_20$ for 60 minutes at room temperature. During this time, a 1 in 3 dilution of Folin & Ciocalteu's phenol reagent was prepared and 0.2ml of the reagent added to each sample tube at the end of the incubation period. The samples were then incubated at room temperature for a further 25 minutes and the OD_{750} measured in a spectrometer. Two controls were used in the assay. In these controls, the sample was replaced with either 25µl or 50µl of a 2mg/ml solution of bovine serum albumin (BSA). The amount of protein, expressed as mg / ml, was then determined using the following calculation:

$O.D._{750}$ sample* volume of standard * 2 ÷ $O.D._{750}$ standard * volume of test assayed

2.2.6 Staining of outer membrane extracts for proteins and LPS

Following SDS-PAGE analysis, outer membrane proteins were stained using the Coomassie staining method (Hancock and Poxton, 1988). The gels were initially placed overnight into solution 1 (1g Coomassie brilliant blue R-250, 500ml propanol-2-ol, 200ml acetic acid, 1300ml distilled water) and then into solution 2 (100mg Coomassie brilliant blue R-250, 200ml propanol-2-ol, 200ml acetic acid, 1600ml distilled water), solution 3 (48mg Coomassie brilliant blue R-250, 200ml acetic acid, 1800ml distilled water) and solution 4 (800ml methanol, 200ml acetic acid, 1000ml distilled water) for 1 hour intervals. Between each stain, the SDS-PAGE gel was washed 3 times in water for 10 minute intervals. The gel was finally placed into solution 5 (200ml acetic acid, 1800ml distilled water) and left in this solution until further use.

SDS-PAGE gels were stained for LPS using the silver staining method (Tsai and Frasch, 1982). The gel was placed overnight into LPS fixative (25% propan-2-ol, 7% acetic acid) and then placed into oxidising solution (0.7% periodic acid / 25% propan-2-ol) for 15 minutes. The gel was subsequently washed 4 times with distilled water for 1 hour intervals. Ammoniacal silver nitrate (0.36% NaOH, 19.4% silver nitrate) was then applied for 15 minutes and the gel was once again washed with distilled water for a 10 minute period. The developer was subsequently added (0.005% citric acid in 0.019% formaldehyde solution) until the gel was sufficiently stained and then washed away with distilled water.

2.2.7 Protein identification by mass spectrometric fingerprinting

Proteins were excised from the gel and reduced, alkylated, and digested as described in section 2.2.3. Peptide-mass fingerprinting subsequently identified them. Spectra were de-isotoped prior to being submitted to the Matrix Science database so that only a single peptide mass was entered into the database for each peptide species. The masses were internally calibrated using the two trypsin autolysis peaks (842.51 and 2211.1) present in the spectra. The calibrated masses of tryptic peptides were then submitted into the Matrix Science database for protein identification. The search criteria restricted the search to the *E. coli* genus and allowed a mass deviation of 50 ppm, oxidation of methionine residues and up to 2 missed cleavage sites by the

enzyme. The probability that the matching protein candidates did not match by chance to the unknown protein was calculated by the database using the MOWSE probability algorithm (Pappin *et al.*, 1993), with scores over 57 being termed a significant hit.

2.2.8 Identification of serum antibodies binding to outer membrane proteins and LPS

E. coli outer membrane extracts were transferred overnight onto nitrocellulose using a BioRad Western blotting tank (5V; 40mA). The following morning, the nitrocellulose was washed in TBS (0.02 M Tris-HCl, 0.5M NaCl; pH 7.5) for 10 minutes and then incubated with 'blocking' solution (3% fish gelatine dissolved in TBS) for 1 hour. The nitrocellulose was subsequently cut into strips and incubated for 3 hours with one of the serum samples diluted 1 in 30 with 'antibody' solution (1% fish gelatine dissolved in TBS). A control was included in which no serum was added to the nitrocellulose strips. The membranes were subsequently washed with TTBS (0.02M Tris-HCl, 0.5M NaCl, 0.025% Tween 20; pH 7.5) for 10 minutes. The bound antibodies were then detected by incubation of the immunoblots with an anti-human horse radish peroxidase-enzyme IgG conjugate for 1 hour (Sigma) and developed using horse radish peroxidase development solution (solution A [60mg 4-chloro-1-naphthol, 20ml methanol], solution B [60µl H2O2, 100ml TBS]; mix together prior to use).

2.3 Methods – part 2

2.3.1 Development of a nano-flow liquid chromatography (LC) system

A nano-flow LC system was set up in the laboratory for the purpose of separating picomole amounts of peptides in complex mixtures. See chapter 4 for a detailed explanation of the system set up.

2.3.2 Sensitivity of the nano-flow LC system

The sensitivity of the nano-flow LC system was tested using calibration mixture 2 (Applied Biosystems) which contained the following synthetic peptides at

different concentrations in the final mixture: Angiotensin I (2.0 pmole/µl), ACTH clip 1-17 (2.0 pmole/µl), ACTH clip 18-39 (1.5 pmole/µl), ACTH clip 7-38 (3.0 pmole/µl), and bovine insulin (3.5pmole/ µl). Different amounts of the synthetic peptide mixture, was injected into the nano-flow LC system to test system sensitivity.

2.3.3 Separation capacity of the nano-flow LC system

The separation capacity of the nano-flow LC system was tested using a tryptic digest of *E. coli* outer membrane proteins. Ten micrograms of *E. coli* outer membrane extracts was resuspended in 100µl of pyrogen free water and boiled in 2% SDS for 5 minutes, reduced at 56°C with 10mM DDT in 100mM $NH_4(HCO_3)_2$ / 5% ACN for 45 minutes, and then alkylated for 2 hours in the dark with 50mM iodoacetamide dissolved in 100mM $NH_4(HCO_3)_2$. Twenty micrograms of trypsin was then added in a tris-phosphate buffer (pH 7.4) and digestion permitted for 4 hour at 37°C. In order to test sensitivity of mass spectrometric detection, the digestion mixture was then diluted 1 in 10^3, 1 in 10^6 and 1 in 10^9 using phosphate buffered saline (PBS). The original digest and diluted digestion mixtures were then passed through strong cation exchange centrifugal devices (Amicon, Millipore) to remove any detergent and the bound peptides subsequently eluted by the addition of 100µl of 1.4N NH_4OH followed by 100µl of 1% TFA. The digestion products were then placed in a speed vac until a volume of 10µl and subsequently diluted 1 in 10 with 0.1% TFA in pure water (ROMIL). To visualise the complexity of the unseparated solution, 10µl of digestion products were concentrated using a Zip Tip™ and analysed by MALDI-TOF MS.

2.3.3.1 One-dimensional and 2 dimensional nano-flow LC separation

Forty five microlitres of a tryptic digest of the outer membrane extracts (~4.5µg outer membrane proteins) were placed into a speed vac to a volume of 10µl and then mixed with reverse phase (RP) buffer A (10% ACN / 0.08% TFA). The extracts were then separated using a C18 15cm x 75µm reverse phase column (LC packings). The method is shown in figure 2.1. During 2D separation, 45µl of a tryptic digest of the outer membrane extracts was dried under vacuum to a volume of

10µl in a speed vac and then diluted 1 in 5 with ion exchange (IEX) chromatography buffer A (10% ACN, 0.08% TFA). The extracts were initially separated using a Magic Bullet strong cation exchange column (Michrom biosciences, USA) and the peptides eluted from the column using 'salt plugs' of NH_4OH of increasing concentration: 50mM, 100mM, 250mM, 500mM, 1M. The peptides eluted by these salt plugs were caught by a C18 capillary trap (LC packings) and subsequently separated by a C18 capillary reverse phase column. For both 1D RP and 2D IEX-RP separation, eluted peptides were mixed 1:1 with saturated matrix solution and deposited directly onto a 400 well MALDI-TOF MS plate ready for analysis.

a)

Time after start event (minutes)	Sequence of events
Start event	Autosampler injects sample
0	Pump starts ACN gradient
15	Computer starts data acquisition
16	Fraction collector begins (collects 60 * 1 minute fractions)
90	Run ends Computer saves data to file

b)

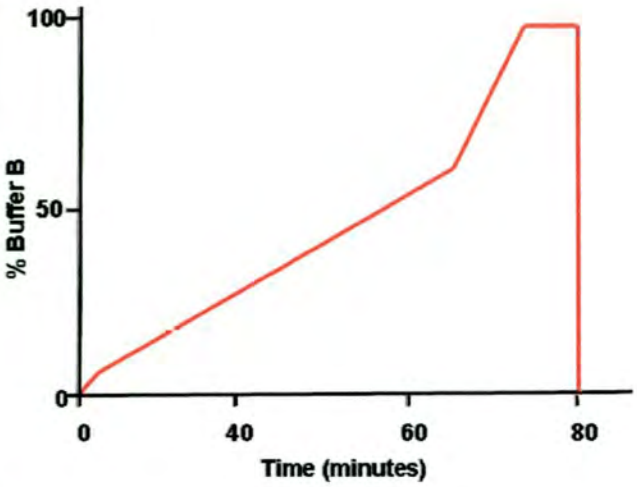


Figure 2.1 Sequence of events and gradient profile used during a reverse phase (RP) chromatographic separation

a) Table showing the computerised events of a RP liquid chromatography run. The different components of the liquid chromatography system were started and coordinated by a computer program package called DAQ factory. The timing of each event is shown in the table.

b) Graph showing the ACN gradient profile used during RP separation. An increasing gradient of ACN was used to elute peptides from the RP column. The gradient reaches 60% buffer B over a period of 65 minutes. The gradient was then rapidly increased to 100% buffer B for 10 minutes and subsequently equilibrated with 100% buffer A for a further 10 minutes.

2.4 Methods – part 3

2.4.1 Production of carbon-13 labelled tetanus toxin C fragment (TTCF)

A histidine-tagged domain of tetanus toxin was expressed and purified as previously described (Hewitt *et al.*, 1997). Recombinant-TTCF includes residues 872-1,315 of the toxin preceded by MGHGHHHHHHHHHHSSGHIEGRHI. *E. coli* expressing recombinant-TTCF was grown at 37 °C in either 1L of LB broth or M9 minimal medium to an OD_{750} value of between 0.5-0.6 and the expression of TTCF induced using 0.2 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) for a period of 3 hours. Carbon-13 glucose was added to the minimal medium at the time of IPTG induction so that it was incorporated into the protein framework during expression. The amount of carbon-13 glucose used for enrichment was 2.3%, given as a percentage of the total amount of glucose in the growth medium. To obtain this percentage, 70mg of carbon-13 glucose was dissolved in 5 ml of filter sterilised water and then added to the growth media. Following induction, bacteria were harvested by centrifugation for 15 minutes at 10,000g, and the pellet re-suspended in 5ml of 'lysis buffer' (25 mM Imidazole / 500mM NaCl / 100mM Tris-HCl). The cells were subsequently lysed by sonication six times at 1 minute intervals to give the total cell protein (TCP) and 10µl removed for SDS-PAGE analysis using a 12 % acrylamide gel.

2.4.2 Purification of carbon-13 labelled tetanus toxin C fragment (TTCF)

Recombinant TTCF was purified using a His-tag column packed with Ni chelated to iodoacetamide (IAA) 4b Sepharose beads (bed depth 3ml; total column length 12 ml). Prior to the addition of the TCP to the His-Tag column, the beads were washed 3 times using 10ml of water and loaded with $NiCl_2$ (0.1M $NiCl_2$, 10mM Na acetate). The column was once again washed 3 times with 10ml of water followed by a further 3 washes with 'conditioning buffer' (same as sonicating buffer). The TCP was then added to the His-tag column, and the protein allowed to bind for 5 minutes before the post load was collected. The column was then washed 3 times with 3ml conditioning buffer and the protein subsequently eluted using 3ml of 'elution buffer' (500mM Imidazole / 500mM NaCl /100mM Tris-HCl). The

elution procedure was repeated twice and all eluates combined. The purified r-TTCF protein was then dialysed extensively into phosphate buffered saline (PBS) for 48 hours and stored at 4°C until use.

2.4.2 Isolation of lysosomes from B cells

Lysosomes were isolated from EBV-transformed lymphoblastoid human Pala B cells as previously described (Davidson *et al.*, 1990; Manoury *et al.*, 1998). In this technique, a combination of cellular lysis, Percoll based centrifugation, and displacement fractionation is employed. Fractions containing lysosomes were identified by the use of a fluorescent enzyme assay and then concentrated by centrifugation for later use.

2.4.2.1 B cell lysis and fractionation of sub-cellular components

B cells were grown to a density of 10^6-10^8 cells, washed once in 10ml of cell culture media and then twice in 10ml of homogenisation buffer (10mM triethanolamine, 10mM acetic acid, 1mM Na_2 EDTA, 0.25M 60% sucrose [w/v]). After each wash, the cells were recovered by centrifugation at 1,500g for 15 minutes. After the final wash, the cells were re-suspended at 3 x 10^7 cells / ml in homogenisation buffer, to give a total volume of 2-3ml. The cells were then passaged on ice 8-10 times through a ball bearing homogeniser (EMBL, Heidelberg, Germany), with a clearance of 6μm. The broken B cells were resuspended in 2-3ml of homogenisation buffer and centrifuged at 2000rpm for 10 minutes at 4°C to obtain the post nuclear supernatant (PNS). The PNS was slowly pipetted onto 23ml of a 27% Percoll solution (10mM triethanolamine, 10mM HAc, 1mM Na_2 EDTA, 0.25M 60% sucrose, 27% Percoll). To fractionate the sub-cellular components, this solution was centrifuged at 56,800g for 1 hour at 4°C and the centrifuge brake switched off to prevent disruption of the sub-cellular gradient during de-acceleration. The gradient was then fractionated using displacement apparatus set up to collect 25 fractions from the bottom of the Percoll gradient upwards. During fractionation, a thin metal tube attached to a pump was gently inserted to the bottom of the centrifugation tube and the pump set to collect fractions at a speed of 1ml / minute. These 1ml fractions were collected into tubes containing 100μl of 0.5 M citrate buffer (pH 5.5).

2.4.2.2 Determination of lysosomal fractions following sub-cellular separation

Fractions containing lysosomes were determined by the use of a fluorometric enzyme assay to detect the activity of β-hexoseaminidase. This enzyme assay was carried out in a 96-well ELISA plate. To all assay wells, 90µl of homogenisation buffer and 100µl of the enzyme substrate 4-methylumbelliferyl n-acetyl β-D-glucosamide (15.18mg dissolved in 0.1M citric acid (pH 4.5), 10% Triton X-100; covered in foil) were added. Prior to the assay, the light sensitive substrate was placed into a 37°C water bath for 15 minutes to dissolve the substrate. To each of the 25 test wells, 10µl of 1 out of the 25 fractions was then added. To a further 10 control wells, 10µl of homogenisation buffer was added in replacement of the sample. The plate was then wrapped in foil and incubated at 37°C for 15 minutes. Following incubation, the reaction was stopped by the addition of 0.1M glycine and the ELISA plate read in a fluorescent plate reader (Cytofluor 4000® Series Fluorescence Multi-well Plate Reader, Applied Biosystems), with emission and excitation wavelengths set at 460nm and 350nm respectively.

2.4.2.3 Pooling of lysosomal fractions and concentration

Fractions containing lysosomes, as determined by the β-hexoseaminidase enzyme assay, were pooled together and concentrated by centrifugation at 110,000g for 1 hour using a Sorvall OTD centrifuge. The lysosomes form a cloudy white film upon a Percoll pellet that can be recovered by careful pipetting. The recovered cloudy white film is washed in 50mM citrate buffer (pH 5.5) and the lysosomes harvested by centrifugation at 89,000g for 1h using a Sorvall OTD centrifuge. The lysosomes were then resuspended in 100µl of 50mM citrate buffer (pH 5.5) and stored at -70 degrees until use.

2.5 Methods – part 4

2.5.1 Identification of labelled TTCF digestion products (<3KDa) following lysosomal digestion

TTCF was digested using lysosomal extracts isolated from lymphoblastoid Pala B cells, as previously described (Manoury *et al.*, 1998). The small digestion

products (<3KDa) were separated using 2D nano-flow LC and characterised using MALDI-TOF MS.

2.5.1.1 Digestion conditions

To disrupt the lysosomal membrane and thus release lysosomal enzymes, lysosomes were mixed with 0.02% CHAPS and left for 5 minutes on ice. The lysosomes (+ 0.02% CHAPS) were then mixed with TTCF in a ratio of 1 to 5 and incubated at 37°C for a period of several hours to allow digestion. All digestions were carried out in the presence of 50mM citrate buffer to keep a pH optimum of either 5.5 or 4.5 during digestion. Ten millimolar DTT was also included in the digestion mixture to reduce the protein. At the end of the digestion period, the reaction was stopped by the addition of 1M NH_4OH. To analyse large digestion species (>3KDa), a proportion of the digestion mixture was ran on a 10 well Novex® 16% Tricine Gel (Invitrogen) based on the tricine system developed by Schaegger and vonJagow, 1987.

2.5.1.2 Two-dimensional IEX-RP nano-flow LC separation of digestion products

The digestion mixture was initially passed through a 3 KDa centrifugal filter unit (Amicon®, Millipore). The 3KDa centrifugal filter unit was subsequently washed with 5% ACN / 0.1 TFA % followed by 90% ETOH. The digestion products were then placed in a speed vac and dried under vacuum to a volume of 10µl before being re-suspended in 50µl of 'IEX buffer A' (H_20 + 0.5% v/v HAc) ready for IEX chromatography. The sample was injected onto a Magic Bullet strong cation exchange column (Michrom, USA) and eluted using a continuous gradient of buffer B (H_20 + 0.5% v/v HAc + 35% ACN + 500mM KCl). The sequence of events during an IEX run and the gradient profile used for separation are shown in figure 2.2. Forty fractions, each collecting 1-minute worth of flow, were collected manually and subsequently pooled together according to their complexity as determined by UV detection. This resulted in a total of 5 IEX fractions that were further separated using RP chromatography. The acid washes were then added to the appropriate fraction. The 5 IEX fractions were prepared for RP separation by drying them in a speed vac to a volume of 10µl and diluting them 1 in 5 using RP buffer A (3% ACN, 0.8%

TFA). The peptides were then injected onto a C18 capillary trap and subsequently separated using a C18 reverse phase capillary column. Eluted peptides were mixed online with matrix solution and deposited upon a 400 well MALDI-MS plate.

a)

Time after start event (minutes)	Sequence of events
Start event	Autosampler injects sample
0	Pump starts KCl gradient
	Computer starts data acquisition
1	Fraction collection starts (40 *1 minute fractions)
50	Run ends Computer saves data to file

b)

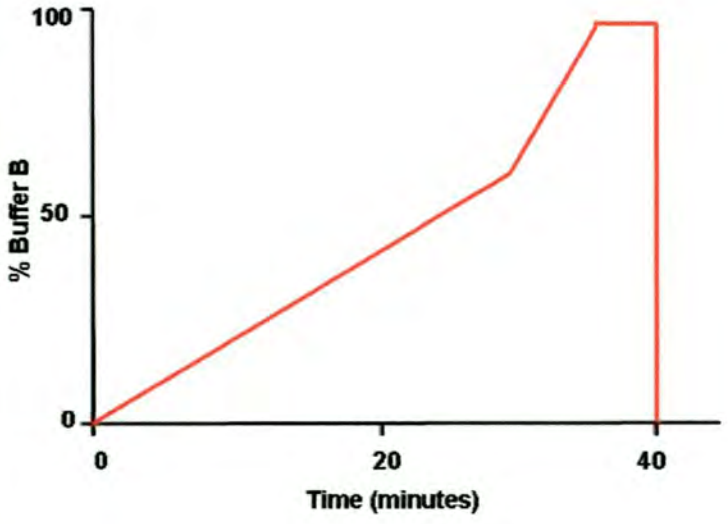


Figure 2.2 Sequence of events and gradient profile used during an ion exchange (IEX) chromatographic separation

a) Table showing the computerised events during a IEX liquid chromatography separation. The different components of the liquid chromatography system were started and coordinated by a computer program package called DAQ factory. The timing of each event is shown in the table.

b) Graph showing the KCl gradient profile used during ion exchange chromatography. An increasing gradient of KCl was used to elute peptides from the strong ion exchange column. The gradient reaches 60% buffer B over a period of 30 minutes. The gradient was then rapidly increased to 100% buffer B over the next 5 minutes and equilibrated with 100% buffer A for a further 10 minutes.

2.5.1.3 MALDI-TOF MS analysis

The set up of the mass spectrometer workstation used in this project is shown in figure 2.3. See section 2.2.4 for sample preparation and data collection. Spectra were calibrated using standard calibration mixture 2, either internally or by the use of a close external calibration. Carbon-13 labelled peptides were identified in spectra by visual analysis or by the use of the in house computer program called TSpec. Initial analysis employed the Data Explorer software package (Perspective). Estimation of the carbon-13 content was made by visual inspection of the isotopic pattern of individual peptides within Data Explorer or by peak analysis with TSpec. The TSpec program was used in conjunction with visual analysis and grouped together peaks likely to constitute isotopes of individual peptides by recognising their compatible mass to charge ratios and relative abundance. The TSpec program also estimated the 'labelledness' of the isotopic patterns for each peptide by calculating two parameters (Zou *et al.* 2004). The first parameter, paramC13%, is an estimate of the carbon-13 content of a peptide calculated from the relative heights of its isotopic peaks. The second, paramλ, is an estimation of the number of carbons per Dalton molecular weight of a peptide calculated from the relative heights of the first two or three isotopic peaks. Using these two parameters, the TSpec program awards a value of 'labelledness' to all of the peptides present in the mass spectra.

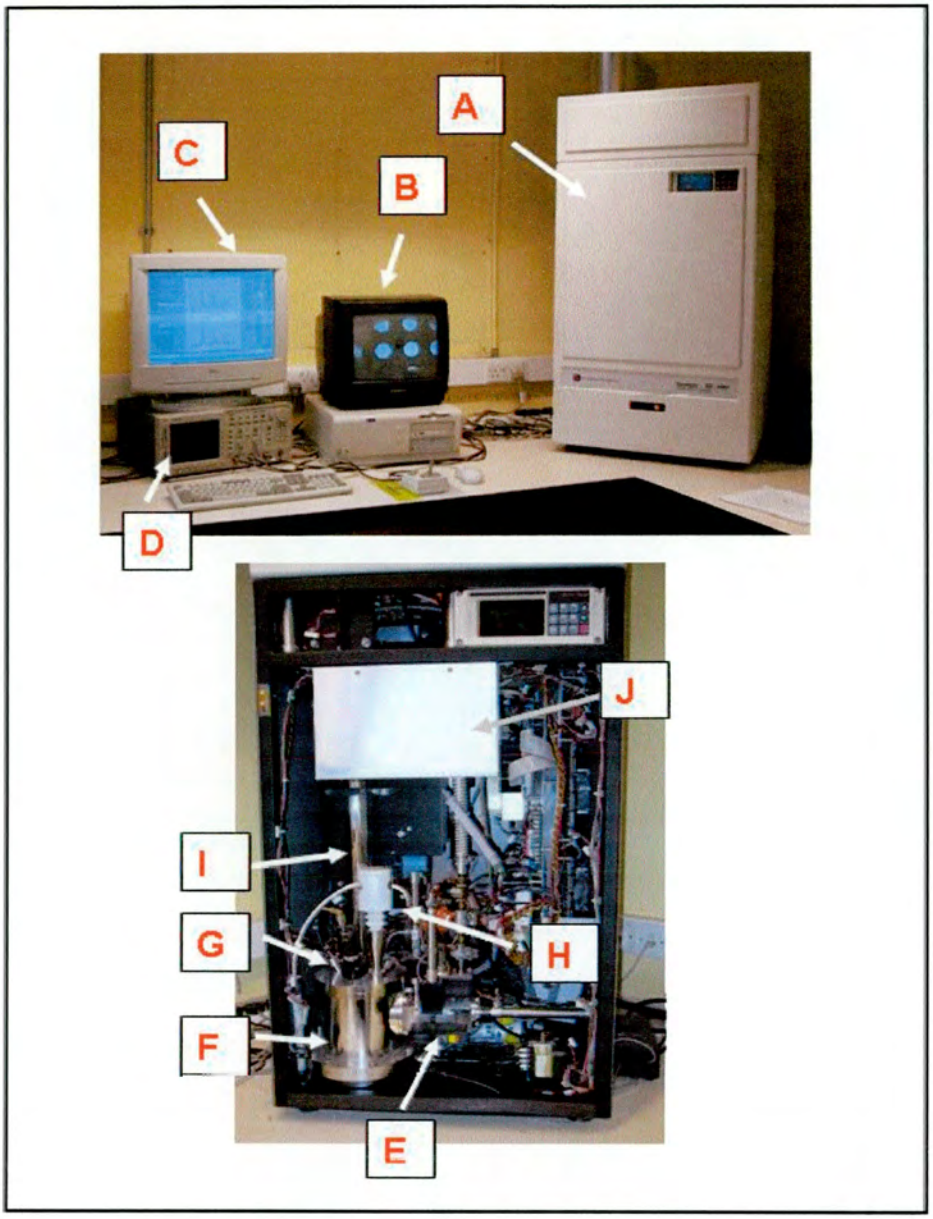

A
C
B
D
J
I
G
H
F
E

Figure 2.3 Photo of the different components of the DE-PRO mass spectrometer workstation
The mass spectrometer DE-PRO workstation consists of four components: (i) a mass spectrometer (ii) a monitor (iii) a computer and (iv) a oscilloscope.

The sample plate is placed into a loading chamber. Upon loading, the sample plate is taken to a main source chamber where the sample is ionised by a laser. The ions are accelerated towards the detector along a flight vessel at a speed determined by the mass of the ion, with lighter ions reaching the detector first. The detector records the signal as an electron cascade displayed on an oscilloscope and shows the data as mass / charge (m/z) against number of ion hits (y axis). Since charge (z) is equal to +1 then m/z is equivalent to the mass of the ion.

Legend
A: Mass spectrometer **B:** monitor **C:** computer **D:** oscilloscope **E:** Sample loading chamber **F:** Main source chamber **G:** camera **H:** laser **I:** Flight tube **J**: Detector

2.6 Methods – part 5

2.6.1 Identification of pre-digested TTCF tryptic peptides in complex mixtures

2.6.1.1 Characterisation of TTCF peptides

Labelled TTCF peptides were characterised following trypsin digestion by mass matching to the TTCF sequence. The sensitivity and specificity of labelling were calculated as followed. The sensitivity was calculated by determining the proportion of TTCF derived peptides that had a labelled isotopic appearance. The specificity was determined by calculating how many peptides not matching to the TTCF sequence displayed an unlabelled isotopic appearance.

2.6.1.2 Identification of pre-characterised labelled TTCF peptides in complex peptide mixtures

The characterised and labelled TTCF tryptic peptides were mixed with unlabelled peptides in two mixtures of differing complexity. In both mixtures, the TTCF peptides were at a relative low abundance. The number of labelled TTCF peptides detectable in both of these complex mixtures was determined before and

after 1D nano-flow liquid chromatography. For the first mixture, 10μg of three unlabelled proteins were digested using trypsin: bovine serum albumin, bovine haemoglobin, and chicken egg lysozyme. Fifty picomoles of this unlabelled digest were mixed with ten pmoles of a tryptic digest of labelled TTCF. This mixture of labelled and unlabelled tryptic peptides was Zip tipped and subsequently analysed using MALDI-TOF MS.

To generate a more complex mixture, ~10μg lysosomes (+5mM DTT / 50mM citrate buffer/ 0.02% CHAPS) were incubated at 37°C for 4 hours to allow autolysis. In order to test sensitivity of mass spectrometric detection, the digestion mixture was diluted 1 in 10^3, 1 in 10^6 and 1 in 10^9 using phosphate buffered saline (PBS) and 10μl mixed with 10 pmoles of a tryptic digest of TTCF. The autolysis products and labelled TTCF peptides were then passed through a 3 KDa cut off membrane. The peptides were washed through the membrane using 50μl of 5% ACN / 0.1% TFA followed by 50μl 90% EtOH, dried to a volume of 10μl in a speed vac (Savant AES 1000/ AES 1010) and then diluted 1 in 5 with RP buffer A.

2.6.1.3 Nano-flow LC separation of complex mixtures

Both the complex mixtures were separated using 1D RP nano-flow LC chromatography. Labelled peptides were subsequently identified by visual inspection of the spectra for the pre-characterised tryptic fragments. A control was included in which 10 pmoles of TTCF tryptic peptides alone was separated using 1D RP nano-flow LC. To test whether familiarity of the tryptic TTCF peptides influenced identification, a person unfamiliar with the characterised tryptic TTCF fragments was also asked to identify labelled peptides in the spectra.

Chapter 3 Identification and carbon-13 labelling of *E. coli* outer membrane proteins

3.1 Introduction

The bacterial antigens chosen for use in this study were outer membrane proteins isolated from *E. coli*. The study chose to focus initially upon outer membrane proteins because the outer membrane represents the barrier between bacteria and their environment. Outer membrane proteins are therefore accessible for antibody binding and are thus favourable targets for immune attack in the clearance of extracellular bacterial pathogens. In addition to their surface exposure, outer membrane proteins are prevalent amongst many different *E. coli* isolates (Overbeeke and Lugtenberg, 1980), and have been shown to elicit protective immune responses (Russo, McFadden *et al.*, 2003). Therefore, the identification of conserved T cell epitopes within outer membrane proteins may prove beneficial in the development of subunit vaccines.

3.2 Objectives

- To show that the outer membrane proteins contain B cell epitopes by examining whether healthy individuals have serum antibodies.
- To identify outer membrane proteins bound by serum antibodies using the mass spectrometry-based technique called peptide mass fingerprinting.
- To generate and purify carbon-13 labelled bacterial outer membrane proteins and to determine the isotope signature of carbon-13 labelled peptides derived from these proteins.

3.3 Results

3.3.1 Outer membrane protein and LPS profiles of different *E. coli* strains

Outer membrane extracts were isolated from three strains of *E. coli* using the well established 'Sarkosyl' technique as described in the methods section. In this procedure, the inner membrane and other proteins and lipids are solubilised by the

detergent Sarkosyl and the outer membrane is subsequently recovered as an insoluble pellet by centrifugation. To examine the protein profile of the *E. coli* K12 outer membrane, the outer membrane extracts were separated by Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and stained with Coomassie blue. Figure 3.1a shows that there are a number of dominantly expressed outer membrane proteins in the outer membrane of *E. coli* K12. Three of these are in the higher molecular weight range of between ~33 and 37 KDa (bands 1, 2, 3) and two are below 20 KDa in molecular weight (bands 4 and 5). These dominantly expressed outer membrane proteins were also present in the outer membrane extracts of the other two strains of *E. coli*. (3.1b). Analysis of the outer membrane extracts by PAGE also revealed the presence of a number of less abundantly expressed proteins which differed considerably between the three strains of *E. coli*.

Outer membrane extracts from two of the *E. coli* strains, K12 (rough LPS) and O18 (smooth LPS) were silver stained in order to detect the presence and phenotype of LPS within the outer membrane extracts (3.1 c & d). LPS is the only lipid in the outer leaflet of the outer membrane and consists of three regions: Lipid A, the core structure, and the O antigen. Examination of the outer membrane extracts isolated from *E. coli* K12 revealed a dense staining at the gel front, corresponding to Lipid A plus core. This is the only part of the LPS molecule found in rough strains of *E. coli* such as K12. In contrast, when examining the extracts isolated from *E. coli* O18, there was additional staining giving rise to a high molecular weight laddered appearance that was absent from the rough strain. This staining corresponds to the O antigen plus core plus Lipid A which is only found in smooth strains of *E. coli*, with each rung of the ladder differing in molecular weight by one repeating unit of O antigen.

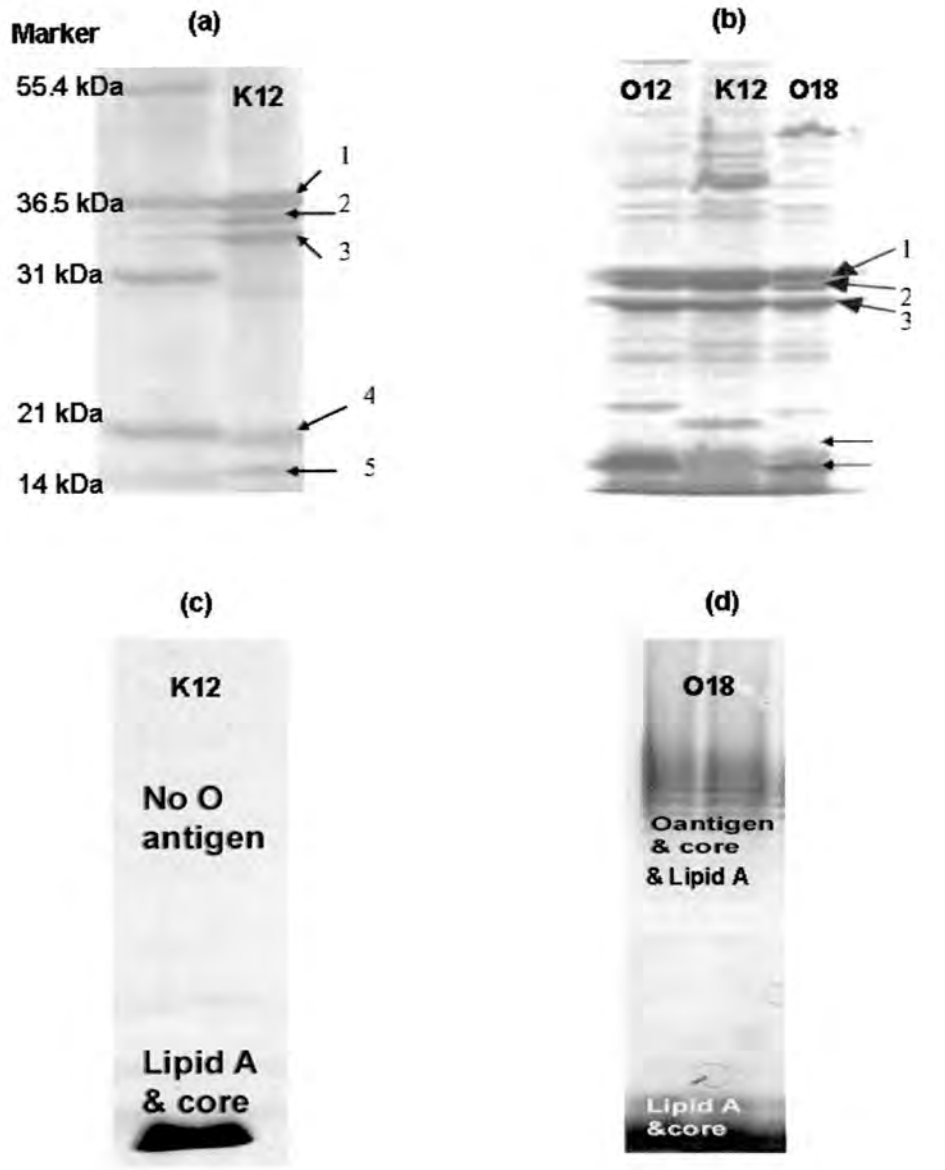


Figure 3.1 SDS-PAGE analysis of the outer membrane protein and LPS profiles of *E. coli*

Outer membrane extracts were separated using a 10% SDS-PAGE gel. **(a)** Coomassie blue staining of the outer membrane extracts of *E. coli* K12 **(b)** The outer membrane protein profile of *E. coli* K12 is compared to *E. coli* O12 and O18. **(c) & (d)** The outer membrane extracts from *E. coli* K12 and *E. coli* O18 have been silver stained to detect the presence of LPS. Arrows 1-5 denotes major outer membrane proteins. Protein bands 4 & 5 are not clear in gel **(b)**.

3.3.2 Detection of outer membrane protein and LPS specific antibodies in healthy individuals

Unstained SDS-polyacrylamide gels of outer membrane extracts from *E. coli* K12 and O18 were transferred onto nitrocellulose and incubated with serum collected from healthy individuals. The immunoblots of *E. coli* K12 show that all 13 individuals had serum antibodies that bound to one or more of the major outer membrane proteins (figure 3.2). However, the levels and pattern of binding differed considerably between the individuals. In most individuals, the serum antibodies bound most strongly to protein band 3 with a molecular weight of approximately ~33 KDa. In comparison, weaker reactions were observed with the other outer membrane proteins. Immunoblots of the *E. coli* O18 extracts using serum from 10 healthy individuals (figure 3.3) showed similar results to *E. coli* K12. However, strong reactions were also observed with protein band 4 (~17KDa), and in some individuals with protein protein band 1 (~37KDa). It is unclear whether band 4 is a low molecular weight protein or whether it is the rough form of LPS (lipid A plus core) that is located at the gel front. In addition, there was also staining corresponding to the binding of serum antibodies to the LPS O antigen. This gave rise to the characteristic ladder pattern superimposed upon the outer membrane proteins in the higher molecular weight range.

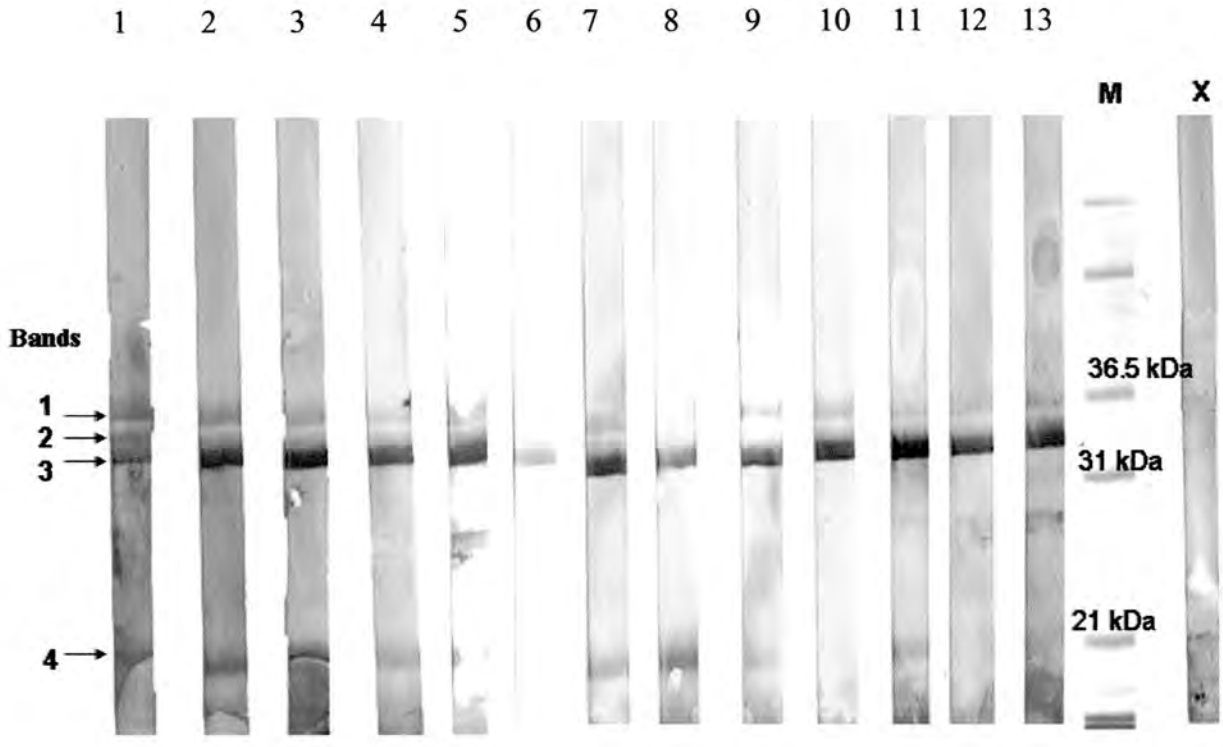


Figure 3.2 Immunoblots showing the presence of serum antibodies that bind to *E. coli* K12 outer membrane proteins from healthy individuals

Serum samples from 13 healthy individuals were immunoblotted against *E. coli* K12 outer membrane extracts. Numbers next to bands indicate outer membrane proteins bound by serum antibodies.

X: No serum control
M: protein marker

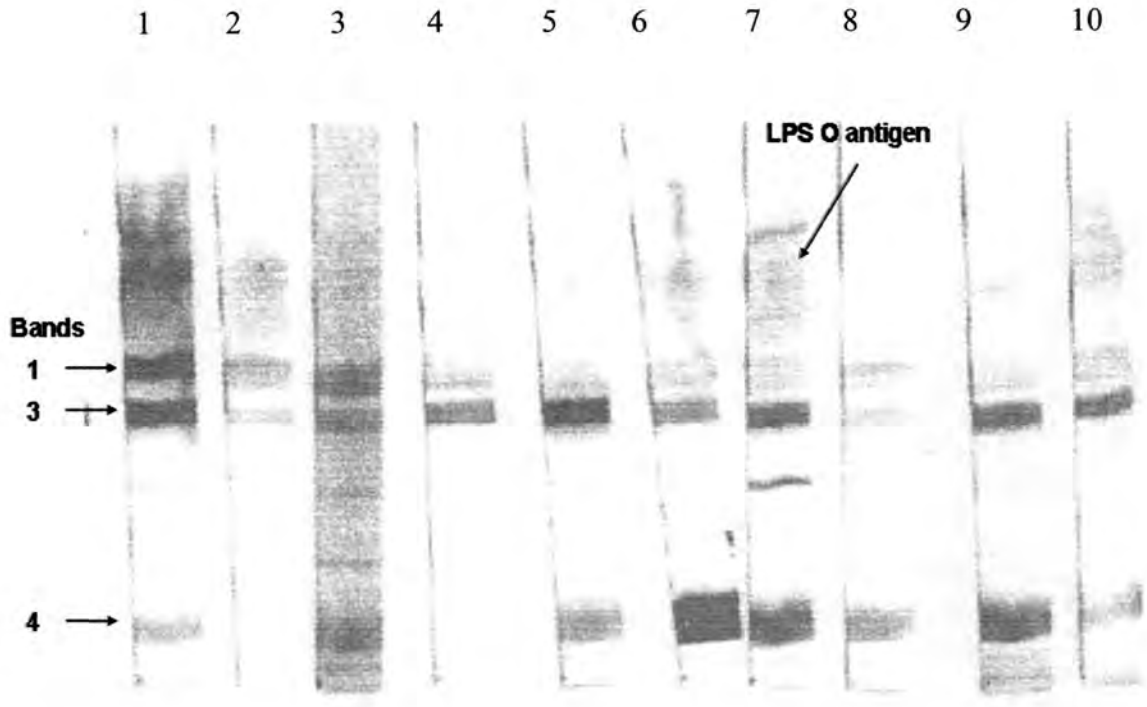


Figure 3.3 Immunoblots showing the presence of serum antibodies binding to *E. coli* O18 outer membrane proteins from healthy individuals

Serum samples from 10 healthy individuals were immunoblotted against *E. coli* O18 outer membrane extracts. Numbers next to bands indicate outer membrane proteins bound by serum antibodies.

3.3.3 Identification of outer membrane proteins

The identities of the three outer membrane proteins within the molecular weight range 30-40 KDa were investigated using the mass spectrometric based technique known as 'peptide fingerprinting' (table 3.1). In this technique, the protein band of interest is excised from the gel and digested with the enzyme trypsin. This generates a characteristic set of peptide fragments that are unique to the protein and are thus termed a 'peptide fingerprint' of the protein. The masses of these peptides are accurately determined by MALDI mass spectrometry and searched against the Matrix Science database. The database contains theoretical trypsin digests of all of the proteins included within the search criteria. The search software compares the theoretical masses of expected fragments of proteins to the experimental masses. Significant protein matches are determined using the MOWSE algorithm (Pappin *et al.*, 1993) and the results returned in the form of a concise summary report. The MOWSE score is a measure of the probability that the experimental masses could be matched to the candidate protein by chance and takes into consideration both the number of peptide matches and the percentage coverage of the candidate protein. MOWSE scores over 57 were considered as significant hits ($p<0.05$).

Using the search criteria stated in the methods section, protein bands 1, 2, and 3 were identified as OMP F, OMP C and OMP A respectively (table 3.1). The database search results for OMP F are shown in figure 3.4 to highlight the method used for identification for all three proteins. In all three cases, the identified proteins were the only proteins in the database likely to account for the major tryptic peptides identified. In addition, sequence coverage of over 50% was observed for all three proteins. The credence that can be placed in these identifications is very high because of the method by which the outer membrane proteins were purified, the high sequence coverage (figure 3.5), the high MOWSE score and by the fact that the matches were of the correct molecular weight.

For each of the three proteins, intense signals in the peptide mass fingerprint matched to expected tryptic fragments of the respective OMP. Peptides likely to derive from OMP C were found in extracts of bands 1 and 3 as well as band 2, presumably as a consequence of contamination during the excision procedure. Unfortunately, band 4 was not successfully identified in this study. Similar studies

have identified a major outer membrane protein of a corresponding molecular weight as peptidoglycan associated lipoprotein (~18KDa). Alternatively, it may be that this band is the rough form of the LPS molecule that would also explain why the band was not identified.

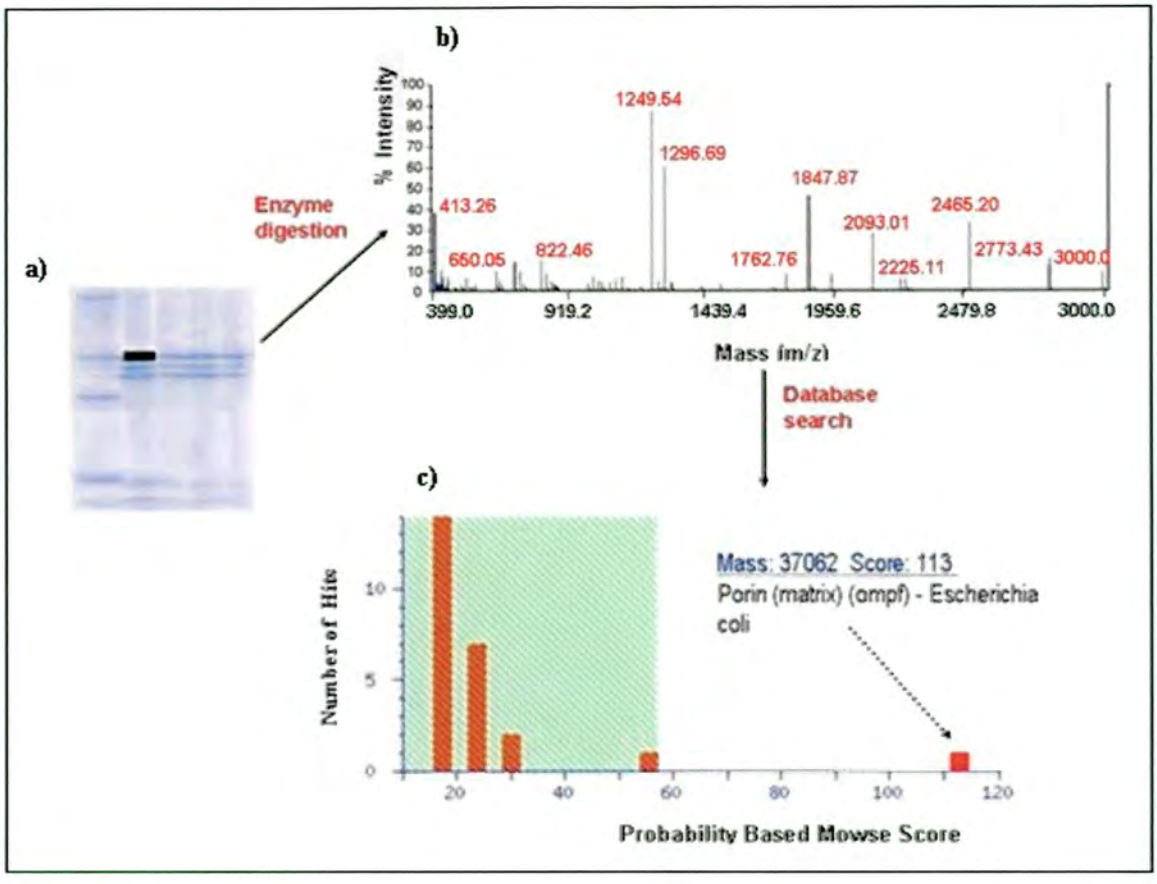


Figure 3.4 Identification of outer membrane protein F

(a) Shown above is the procedure carried out during mass spectrometric identification of proteins. The protein band is excised from an SDS-PAGE gel and digested by trypsin. This generates a characteristic set of peptide fragments that are unique to the protein and are thus termed a 'peptide fingerprint' of the protein. The masses of these peptides are shown in **(b)** and are accurately determined by MALDI mass spectrometry. The calibrated masses are submitted into the Matrix Science database, which carries out theoretical trypsin digests of all of the proteins included within the search criteria. **(c)** Significant protein matches are determined using the MOWSE algorithm and the results returned in the form of a concise summary report. The MOWSE score is a measure of the probability that the experimental masses could be matched to the candidate protein by chance and takes into consideration both the number of peptide matches and the percentage coverage of the candidate protein. MOWSE scores over 57 are considered as significant hits ($p<0.05$).

	Protein identity	Number of peptides submitted	Number of peptides matched	MOWSE score	% Coverage
Band 1	OMP F	50	14	113	50
Band 2	OMP C	53	18	181	66
Band 3	OMP A	39	11	107	55

Table 3.1 Identities of the major outer membrane proteins using peptide mass fingerprinting

The identity of the three major outer membrane proteins within the molecular weight range 30-40 KDa are shown in the table above.

See appendix 2 for more information on the database search results

OMP C

```
  1 MKVKVLSLLV PALLVAGAAN AAEVYNKDGN KLDLYGKVDG LHYFSDDKSV
 51 DGDQTYMRLG FKGETQVTDQ LTGYGQWEYQ IQGNSAENEN NSWTRVAFAG
101 LKFQDVGSFD YGRNYGVVYD VTSWTDVLPE FGGDTYGSDN FMQQRGNGFA
151 TYRNTDFFGL VDGLNFAVQY QGKNGSVSEG MTNNGREALR QNGDGVGGSI
201 TYDYEGFGIG AAVSSSKRTD DQNSPLYIGN GDRAETYTGG LKYDANNIYL
251 AAQYTQTYNA TRVGSLGWAN KAQNFEAVAQ YQFDFGLRPS LAYLQSKGKN
301 LGVINGRNYD DEDILKYVDV GATYYFNKNM STYVDYKINL LDDNQFTRDA
351 GINTDNIVAL GLFYQF<
```

Sequence matches are shown in red

Figure 3.5 Sequence coverage of OMP C

Shown above are the sequences of OMP C matched by experimental peptide masses generated by trypsin digestion of protein band 2.

See appendix 2 for the sequence coverage of OMP A and OMP F

3.3.4 Production of Carbon-13 labelled outer membrane proteins

E. coli O18 was grown in M9 minimal medium enriched with different amounts of carbon-13 glucose. The results in figure 3.6 show that as the amount of carbon-13 supplementation was increased, there was an increase in the number of isotopic peaks observed for each individual peptide species and also a change in the intensity of the isotopic peaks relative to one another. Together, these changes resulted in a change in the overall isotopic pattern of the peptide species. As carbon-13 enrichment was increased from 2.3 to 3.3%, there was a corresponding increase in the contrast of appearance of the isotopic pattern for the unlabelled and labelled peptide. However, at the higher enrichment level of 3.3% carbon-13, there is an observable decrease in sensitivity. This is noticeable with the peptide of m/z value 2977.5 that has disappeared into the baseline of the mass spectrogram at the 3.3% carbon-13 enrichment level. In addition, figure 3.7 shows that at the higher enrichment level, the isotopic envelopes of neighbouring peptides 'overlap' with one another, making them unidentifiable as two labelled peptides. This decrease in sensitivity and occurrence of overlapping peptides at the 3.3% carbon-13 enrichment level was undesirable. As a consequence, enrichment with 2.3% carbon-13 was chosen as the optimal amount of carbon-13 enrichment.

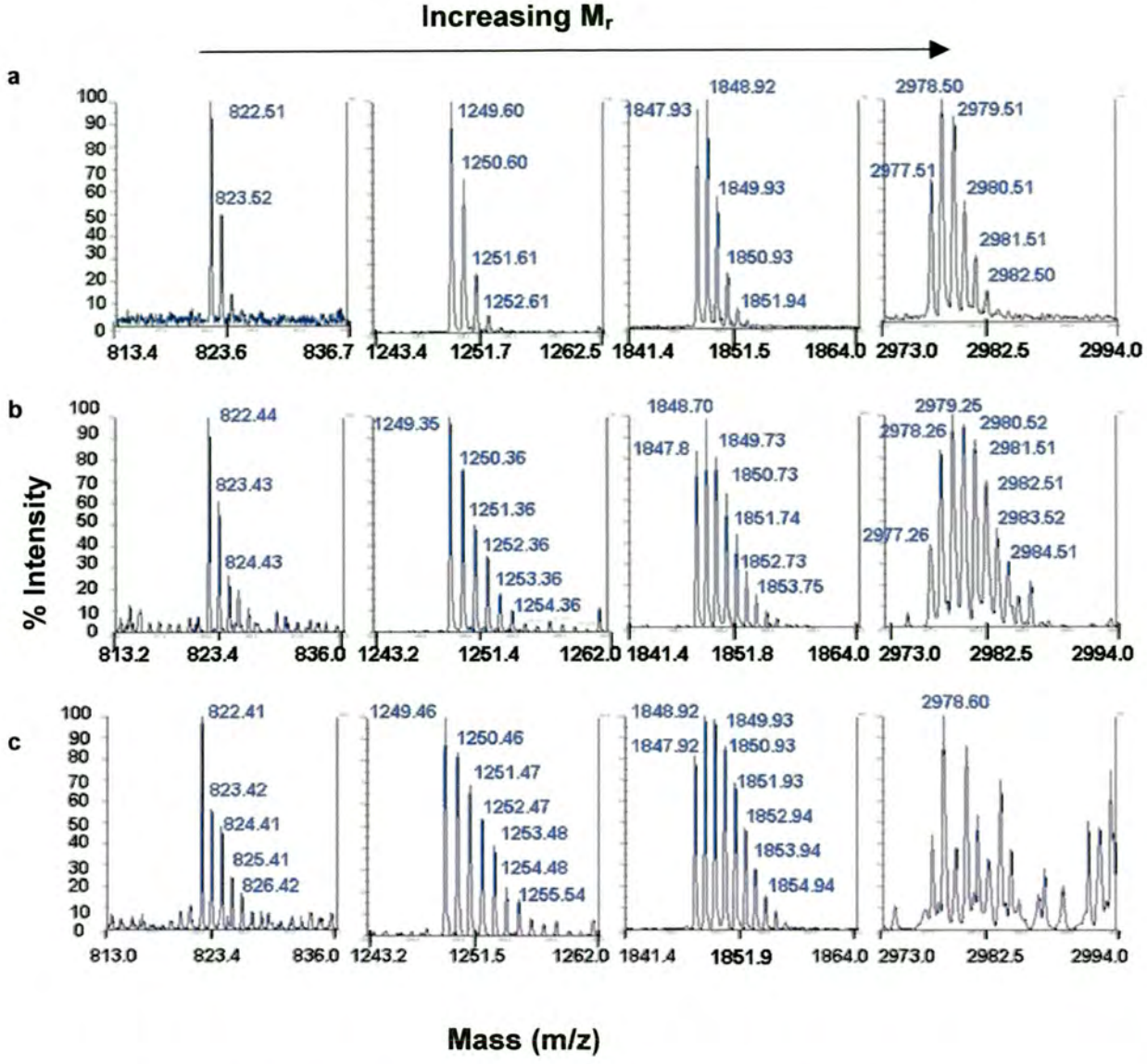


Fig 3.6 Determination of the optimum amount of carbon-13 enrichment of the outer membrane proteins of *E. coli*

E. coli was grown in M9 minimal medium with either **(a)** 1.1% carbon-13 **(b)** 2.3% carbon-13 or **(c)** 3.3% carbon-13. The TTCF protein was then purified using a His-tag purification procedure and the purified protein digested with trypsin. The spectra above show four of the tryptic peptides of differing mass, which are derived from OMP F. For each peptide, the isotopic pattern for the unlabelled peptide is compared to that following enrichment with 2.3% and 3.3% carbon-13.

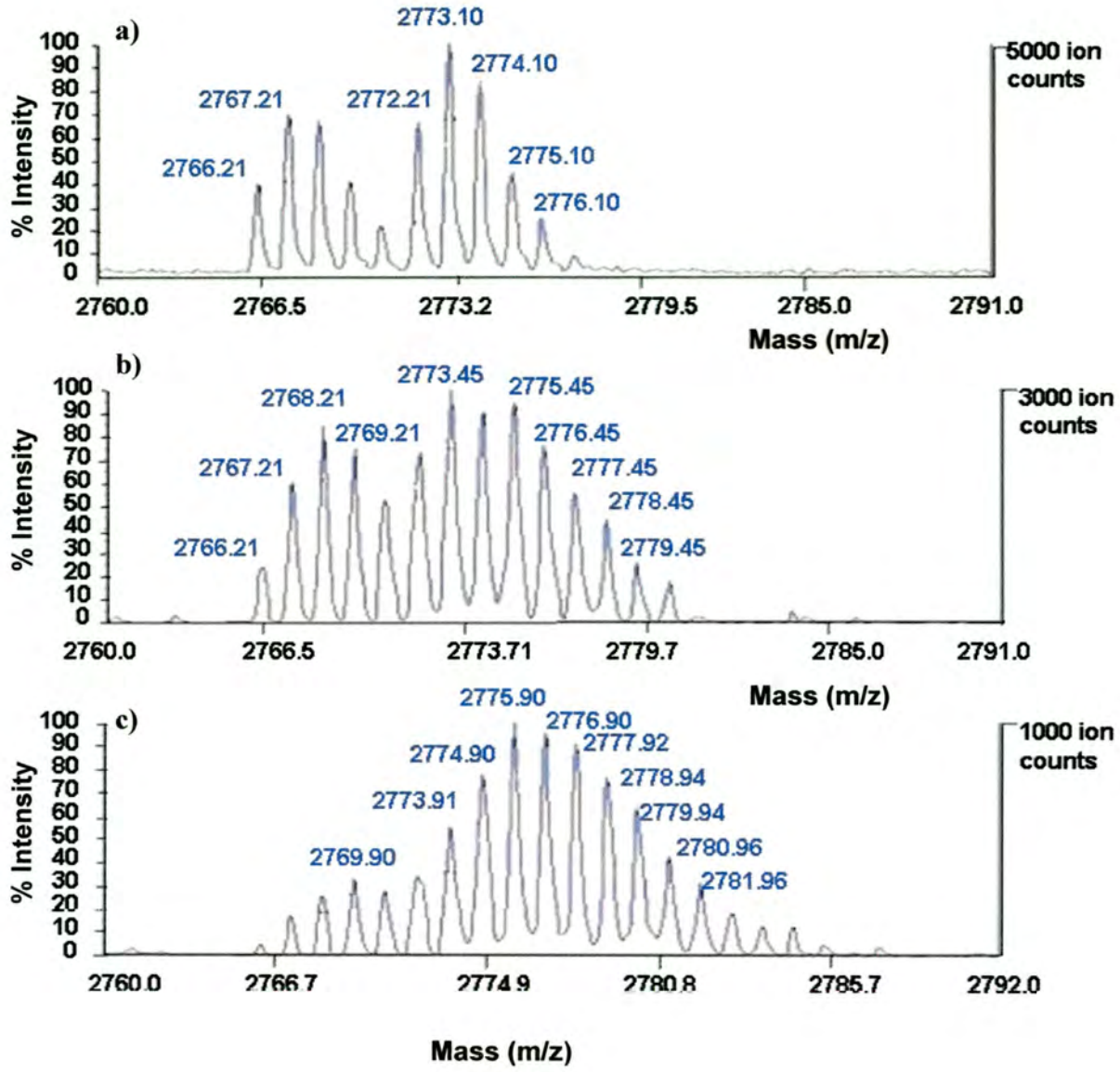


Figure 3.7 Deleterious effect of carbon-13 labelling of proteins

Three mass spectrograms showing the isotopic pattern of two OMP F peptides at three different levels of carbon-13 content which were **a)** 1.1%, **b)** 2.3% and **c)** 3.3%

3.4 Discussion

In the experiments within this chapter, outer membrane extracts were isolated from three strains of *E. coli* and analysed using a combination of SDS-PAGE and immunoblotting to examine their complexity and immunological relevance. These outer membrane extracts consist of a number of proteins, phospholipids and lipopolysaccharide (Benz 1988). *E. coli* was subsequently grown in carbon-13 enriched minimal medium to determine whether peptides derived from labelled bacterial proteins could be distinguished from those derived from unlabelled proteins and to determine the appropriate amount of labelling.

3.4.1 Outer membrane protein profiles of *E. coli*

The results from this study showed that there are a handful of dominantly expressed proteins in the outer membrane of *E. coli* as well as a number of proteins expressed in smaller amounts termed minor outer membrane proteins. It has previously been determined that there are over 60 outer membrane proteins that can be expressed by *E. coli* (Molloy *et al.*, 2000). However, some of these may not have been expressed under the growth conditions used this project and others may be expressed too weakly to be detected using 1D SDS-PAGE. The results from this study showed that the dominantly expressed outer membrane proteins were conserved between the three different strains of *E. coli* that were investigated. This is in accordance with other studies showing that major outer membrane proteins are conserved between different strains of *E. coli* and show similarities to different members of the Enterobacteriaceae family including *Salmonella* and *Shigella* (Hofstra and Dankert 1980; Overbeeke and Lugtenberg 1980). In contrast to the major outer membrane proteins, differences were observed in the minor protein profiles between the three strains of *E. coli*. These differences in the minor outer membrane protein profiles may have resulted from gene mutations, leading to proteins with slightly different sequences and thus molecular weights. Alternatively, they may be different proteins with similar molecular weight values.

The major outer proteins in the molecular weight range 30-40 KDa were identified as OMP F, OMP C and OMP A. These results are in accordance with other studies that have shown the presence of these proteins in the same molecular range.

Although the lower molecular weight proteins were not identified in this study, the literature identifies two proteins of corresponding molecular weight as peptidoglycan associated lipoprotein (17.5KDa) and Braun lipoprotein (7KDa). As mentioned in the results section, it could also be the case that the lower molecular band may be the rough form of the LPS molecule. These major outer membrane proteins play different roles in bacterial physiology and in some cases have been implicated in pathogenesis. Both OMP C and OMP F are general diffusion porins that are constitutively expressed in the outer membrane. These trimeric porin structures within the outer membrane allow the diffusion of nutrients through this otherwise impermeable 'barrier'. Murein lipoprotein, peptidoglycan associated lipoprotein and OMP A are all thought to play an important role in outer membrane stability and have been implicated to play a role in sepsis (Hellman, Loiselle *et al.*, 2000; Hellman and Warren, 2001; Hellman, Roberts *et al.*, 2002; Hellman, Tehan *et al.*, 2003; Liang, Bagchi *et al.*, 2005). In addition, OMP A has been implicated in the pathogenesis of meningitis in *E. coli* K1 and has been shown to be a target of the host defence (Prasadarao, Wass *et al.*, 1996; Prasadarao, Wass *et al.*, 1999; Belaaouaj, Kim *et al.*, 2000; Prasadarao, Blom *et al.*, 2002; Sukumaran and Prasadarao, 2002; Prasadarao, Srivastava *et al.*, 2003; Sukumaran and Prasadarao, 2003; Sukumaran, Shimada *et al.*, 2003; Sukumaran, Selvaraj *et al.*, 2004; Selvaraj and Prasadarao, 2005).

3.4.2 Immunological relevance of outer membrane proteins

The immunoblots revealed that a number of the outer membrane proteins bound to serum antibodies from healthy individuals and thus contained B cell epitopes. However, the nature and intensity of binding varied between individuals. Previous studies have also shown the presence of serum antibodies in healthy humans that bind to *E. coli* surface proteins including OMP A, lipoprotein and iron binding proteins (Griffiths, Yoonessi *et al.*, 1977; Griffiths, Stevenson *et al.*, 1985; Henriksen and Maeland 1987). These antibodies are mainly of the class IgG and have been found in infants of below one years of age. It has been suggested that they are generated against gut bacteria that colonise the intestinal tract shortly after birth (Griffiths, Yoonessi *et al.*, 1977). The results from this study showed that serum

antibodies from the same individuals reacted against outer membrane proteins from different *E. coli* strains. Indeed, the outer membrane proteins from different strains of *E. coli* and between different members of the family Enterbacteriaceae have been shown to be antigenically cross reactive (Hofstra and Dankert 1979; Hofstra, Van Tol *et al.*, 1980; Henriksen and Maeland 1987). This cross reactivity initially raised the possibility that these antigens may be useful in vaccines.

Although the serum antibodies from healthy individuals bind to a number of the major outer membrane proteins, the strongest reactions were consistently observed with OMP A and a protein of ~20 KDa thought to be peptidoglycan lipoprotein. Serum antibodies appeared to react less strongly against the porin proteins OMP F and OMP C. This may be due to more antibodies or stronger avidity antibodies reacting with these two proteins. However, it may have also been due to the fact that the outer membrane preparations were boiled in 2% SDS for 5 minutes prior to PAGE. Boiling in SDS has previously been shown to denature porin trimers into their monomeric forms. This conformational change results in antibodies that are generated against the intact trimer binding only weakly to the denatured monomer (Hofstra and Dankert 1979).

It was a concern in this study that antibodies binding to the outer membrane proteins may be in fact be binding to the LPS component of tightly associated protein–LPS complexes. The results showed that the outer membrane extracts did contain LPS as expected and that serum antibodies from the individuals did bind to the O antigen of the LPS molecule. Other groups have demonstrated that protein bands excised from SDS-PAGE gels are frequently contaminated with LPS following boiling in SDS which is not sufficient to disaggregate the complexes (Chart and Griffiths, 1985; Rocque, Coughlin *et al.*, 1987). Furthermore, it was shown that methods in which detergents are used to recover the outer membrane give rise to protein / LPS complexes which react with serum antibodies (Poxton *et al.*, 1985). These observations mean that one must consider that serum antibodies may be binding to the LPS component of these complexes. Although this was not examined further in this study, similar studies have used different methods to determine the effect of LPS prior to examining the antibody response to outer membrane proteins. This includes destruction of the protein component by treatment with proteinase K

(Nicolle, Ujack *et al.*, 1988; Hellman, Zanzot *et al.*, 1997) or by preincubation of the serum sample to remove anti-LPS antibodies (Bolin and Jensen, 1987; Henriksen and Maeland, 1987). In these studies, it was shown that a proportion of the antibodies did indeed bind to the outer membrane proteins and not LPS.

3.4.3 Production of carbon-13 labelled outer membrane proteins

Glucose is the only carbon source in M9 minimal medium, so the carbon-13 content of bacterial proteins can be easily adjusted by controlling the proportions of natural glucose and carbon-13 glucose in the medium. The results confirmed that a mere doubling in the amount of carbon-13 in the growth medium was sufficient to lead to a detectable alteration in the isotopic pattern of peptides without leading to deleterious effects. This is in line with previous data from the laboratory which has shown that labelling with 2.3% carbon-13 glucose gives optimal labelling for the recombinant proteins sGT and α3(IV)NC1 (Zou, Turner *et al.*, 2004). Although higher amounts of carbon-13 would further enhance the distinction between unlabelled and labelled peptides, 2.3% enrichment was chosen because higher enrichment levels led to the occurrence of two deleterious effects. Firstly, the occurrence of overlapping peptide clusters is increased with an increase in carbon-13 enrichment. This is due to an increase in the number of isotopic peaks for the peptide species that led to an increase in baseline spreading. This is undesirable because it is impossible to distinguish between a labelled peptide and overlapping peptide clusters. In addition, carbon-13 enrichment leads to a decrease in detection sensitivity. This is due to the total number of ion counts being spread over a larger number of isotopic peaks for a given peptide. This is not favourable for the identification of low abundance peptides in complex mixtures that may not be detected at higher carbon-13 enrichment levels due to decreased sensitivity.

Chapter 4 Development of a 1D and 2D nano-capillary liquid chromatography system for the separation of complex peptide mixtures

4.1 Introduction

The work described in the previous chapter indicated that carbon-13 labelling of bacterial antigens was feasible and enormously promising, but revealed a serious shortcoming. It was apparent that labelled peptides are not discernible in overly complex spectra due to the presence of overlapping peptide species. Therefore, the principal aim of this work was to develop a separation technique that allowed the detection and recognition of isotope patterns for peptides at picomole levels in complex mixtures. However, it was a concern that a single dimension separation technique might have insufficient resolution to overcome the problem. In addition, it was also a concern that the ionisation technique proposed to be used, MALDI-MS, could be susceptible to suppression of peptide signals in overly complex spectra. As a result of these two factors, it was reasoned that successful use of carbon-13 labelling within highly complex samples would require a very efficient separation system prior to mass spectrometric analysis. It was anticipated that 1D separation would prove inadequate at such a task and therefore it was envisioned that a 2D system would need to be developed.

The most promising technique at the initiation of this project for resolving and detecting peptides at picomole levels in complex mixtures was 2D nano-flow liquid chromatography. Two-dimensional chromatography involves the sequential application of two orthogonal chromatographic steps, that is to say, separation according to two independent chemical properties of peptides. In the first dimension, ion exchange chromatography (IEX) is employed to separate the peptides based upon cationicity (net positive charge). This is followed in the second dimension by peptides being separated by their hydrophobicity, using reverse phase (RP) chromatography.

4.2 Objectives

• To develop a 2D nano-flow LC system capable of separating complex mixtures of picomole amounts of peptides.

• To determine whether a 2D nano-flow LC system leads to better separation compared to a 1D nano-flow LC system by comparing the number of peptides co-eluting at each retention time, the occurrence of overlapping peptides in individual fractions and the total number of detectable peptides following separation.

4.3 Results

4.3.1 Set-up of the nano-flow LC system

This project initially opted to use a fully automated nano-flow LC system that is shown in figure 4.1a. In this system, peptides were separated using either a single or multidimensional separation technique. During single dimensional chromatography, the peptides were separated by reverse phase chromatography (1D RP) alone. In contrast, during multi-dimensional chromatography, peptides were initially separated by IEX chromatography followed by reverse phase separation (2D IEX-RP) in the second dimension.

During 1D-RP separations, the sample is loaded by an autosampler onto the capillary trap. The capillary trap is then switched in line with the capillary flow path of a C18 (15cm x 75μm) RP column (LC packings, USA) and the peptides subsequently eluted by an increasing gradient of ACN. These eluted peptides pass through a UV detector, are mixed 1:1 with matrix solution and then directly deposited upon a 400 well MALDI-plate for mass spectrometric analysis (figure 1b). A splitter system is used to obtain a flow rate of 200 nl / minute which passes through the RP column. The matrix pump is also set at a speed of 200 nl / minute, collecting a total of 400 nl / minute in each MALDI-MS fraction. During 2D IEX-RP separation (shown in figure 4.2), the peptides are initially injected by an autosampler onto a Magic Bullet ion exchange column (Michrom BioResources, USA). Peptides are subsequently eluted from the column in a stepwise manner using salt 'plugs' of increasing concentration of NH_4OH. Salt-eluted peptides are trapped onto a RP capillary trap (LC packings) which is then switched in line with a capillary RP

column. A 10 port 2-position valve that is shown in figure 4.3 controls the switching of this capillary trap into the flow path of the RP column. The switching of this valve from one position to another changes the flow path of the LC system. Following RP separation, eluted peptides subsequently pass through the UV detector and are deposited upon a 400 well MALDI-plate.

Peptide losses were a major concern, particularly moving from 1D RP to 2D separation. Several modifications were made in an attempt to minimise losses. The system was designed to have a very small wetted area and automated fraction collection was developed so that peptides were deposited directly onto a MALDI-MS plate after elution from the RP column and UV detection.

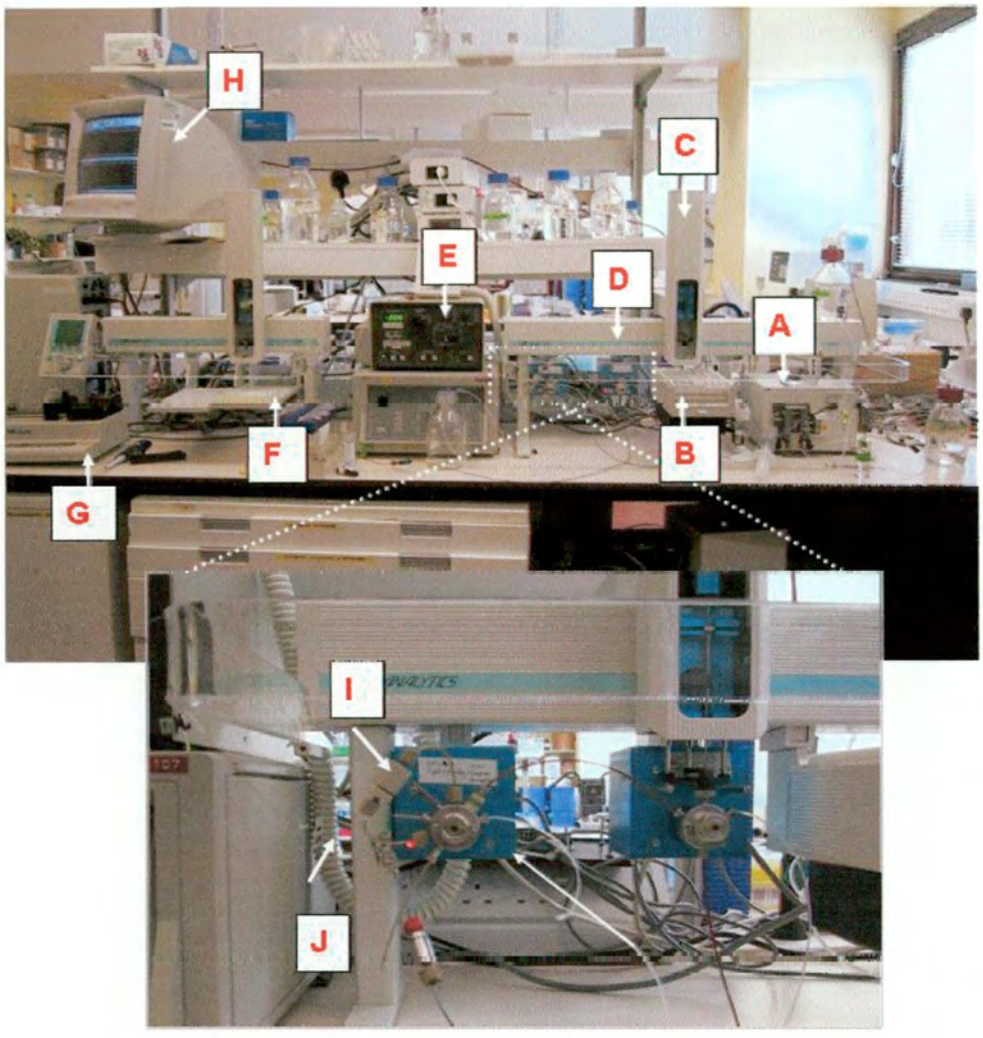


Figure 4.1a. Set-up of the nano-flow RP LC system

Photo showing the different components of the nano-flow LC system set up in the laboratory. The system consists of several 'online' components including a pump, sample stage, autosampler, UV detector, and sample stage for automated sample collection.

Legend
A: HPLC pumps **B:** Sample stage **C:** Autosampler **D:** Injection valve **E:** UV detector **F:** collection stage with MALDI MS plate **G:** matrix pump **H:** computer **I:** Capillary trap **J:** RP capillary column.

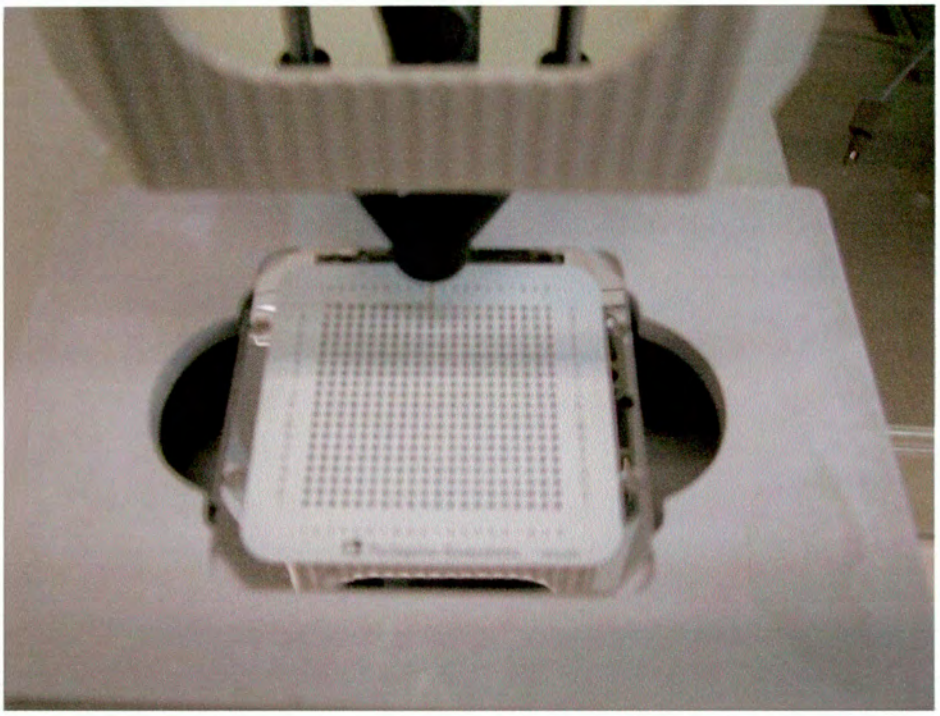


Figure 4.1b Photo showing automated fraction collection onto a 400 well MALDI – MS platc.

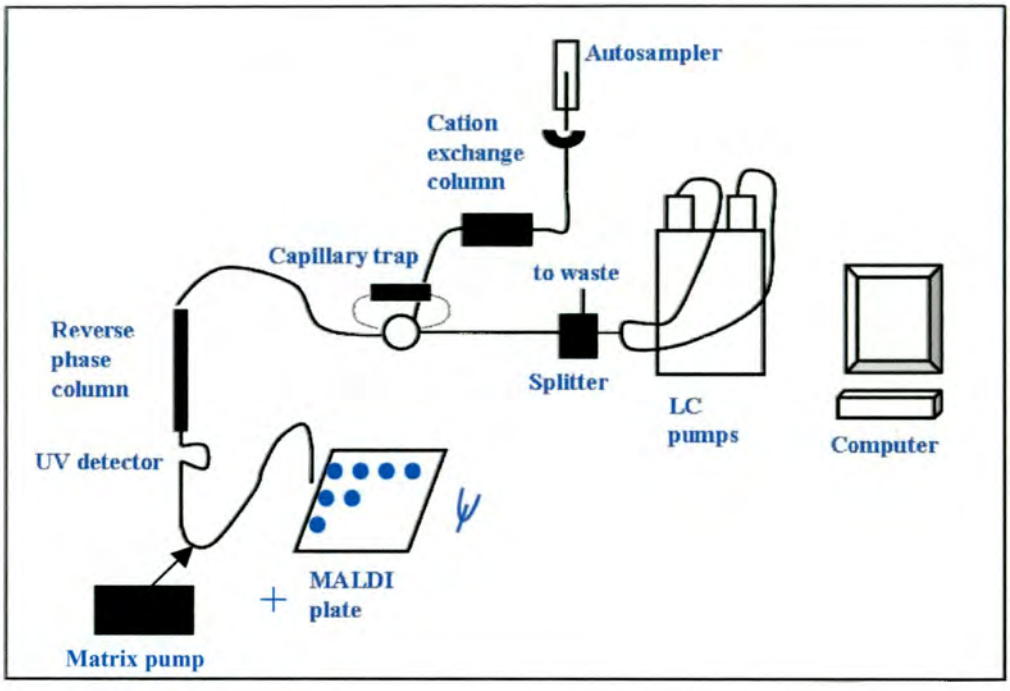


Figure 4.2 Set up of the 2D nano-flow LC system using the 'salt plug' method

The sequences of events are described below:

a) Autosampler injects peptide sample onto the cation exchange (IEX) column followed by a salt plug.

b) Eluted peptides are 'caught' on a capillary trap.

c) Capillary trap switched inline with the C18 capillary RP column by a 10-port 2-position valve and RP separation carried out.

d) Following a RP run, a second 'salt plug' in injected onto the IEX column to elute a different population of peptides and a RP separation is subsequently carried out as described in **(b)** and **(c)**.

e) The method continues until the salt plug of the highest salt concentration has been injected.

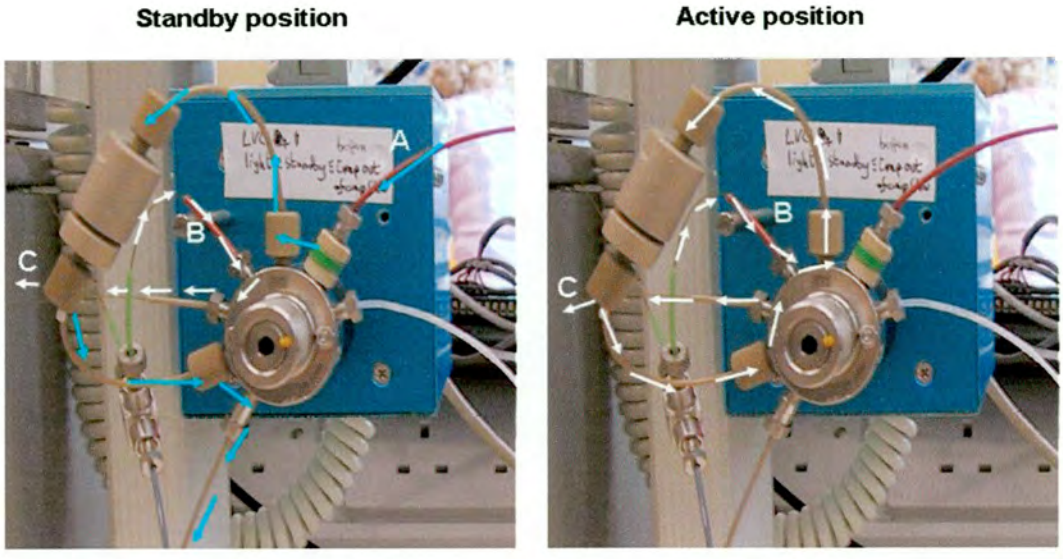


Figure 4.3 Different flow paths of the nano-flow LC system

Shown above are 2 photos of the 10 port, 2 position valve in its different positions. In the standby position of the valve, the capillary trap is out of the ACN flow path. This allows peptides in the sample to be caught by the trap for concentration and desalting. In the active position of the valve, the capillary trap is switched in line with the ACN gradient, allowing peptides to be eluted from the trap and caught upon the RP column for separation.

Legend:
Blue arrows: path of sample
White arrows: path of ACN gradient
A: Position of entry of sample into the valve (from IEX column)
B: Position of entry of ACN gradient into the valve
C: Exit of gradient and sample from the valve (to RP column)

4.3.2 System sensitivity

Once the nano-flow LC system had been set up, the sensitivity of the system was examined using a number of synthetic peptides shown in figure 4.4. The synthetic peptides used in these experiments are described in section 2.3.2 in the methods and materials chapter. Decreasing amounts of synthetic peptides were injected into the nano-flow LC system and separated using reverse phase chromatography. The results showed that the system was highly sensitive, being able to detect 5-10 picomoles of peptides in the mixture. At this level, peptides were also easily detectable by MALDI–MS analysis (results not shown).

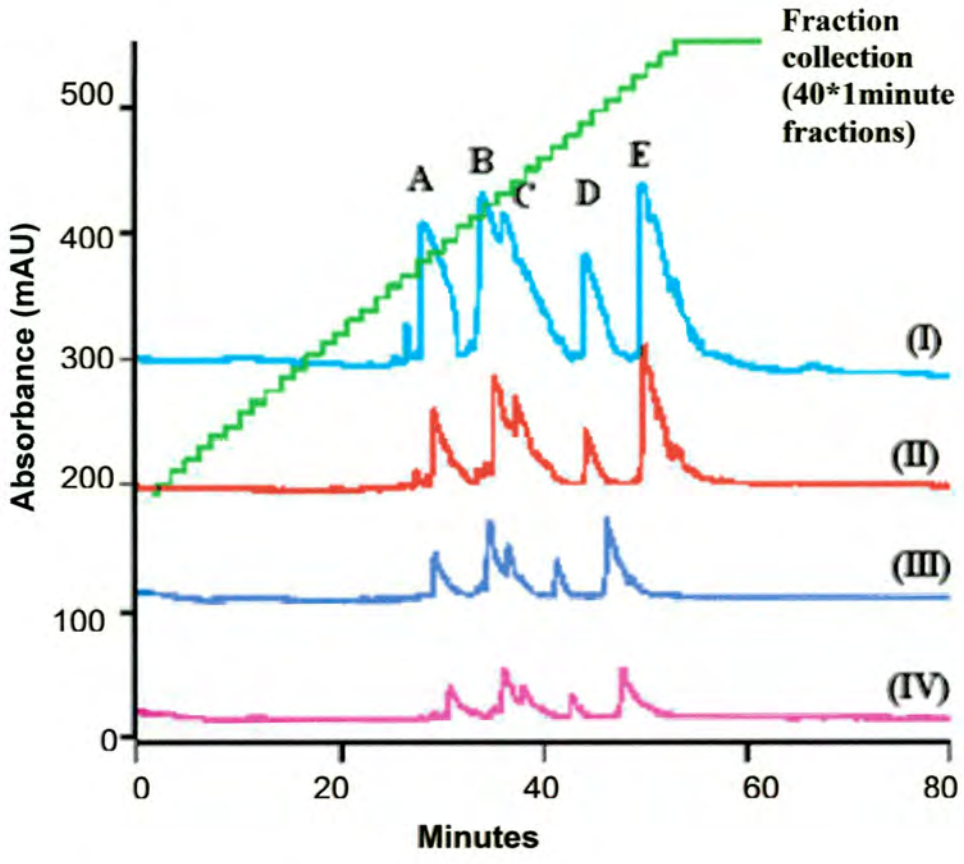


Figure 4.4 Sensitivity of the nano-flow LC system
Absorbance was measured at a wavelength of 214 nm. Different amounts of calibration mixture 2 were injected into the nano-HPLC system to determine whether the system could detect picomole amounts of peptide. Calibration mixture 2 contained the following synthetic peptides at different concentrations in the final mixture: Angiotensin I (2.0 pmole/µl), ACTH clip 1-17 (2.0 pmole/µl), ACTH clip 18-39 (1.5 pmole/µl), ACTH clip 7-38 (3.0 pmole/µl), and bovine insulin (3.5pmole/ µl). The amounts injected were as follows (i) 40 µl, (ii) 20 µl, (iii) 10 µl, (iv) 5 µl.

4.3.3 One and 2D nanoflow liquid chromatography separation of the outer membrane proteins

The goal was to resolve mixtures of labelled and unlabelled peptides, such that the distinctive isotope pattern of labelled peptides was discernable in MALDI mass spectrograms. To achieve this, it was decided that the resolving capability of the nano-flow LC system should be evaluated using the complex peptide mixture obtained by the tryptic digestion of *E. coli* K12 outer membrane proteins which is shown in figure 4.5. The unseparated tryptic digest is extremely complex, with many overlapping peptide species. The experiments were designed to investigate whether the new 2D nano-flow liquid chromatography system could efficiently separate out the tryptic digest of *E. coli* outer membrane proteins, thus preventing the occurrence of overlapping peptide species during mass spectrometric analysis. During 2D IEX-RP separation, the peptides were initially separated into 5 pools by elution from a strong cation exchange column with 'plugs' of increasing salt concentration, and the peptide pools subjected to reverse phase separation. In order to assess the capability of the 2D IEX-RP nano-flow liquid chromatography system compared to 1D RP chromatographic techniques, the tryptic digest of *E. coli* outer membrane proteins was also separated using a 1D RP nano-flow liquid chromatography system, in which the peptides were separated using only RP chromatography. The 1 and 2D separations were compared by determining:

1) The number of peptides co-eluting into fractions collected at equivalent retention times during both a 1D RP separation and a 2D IEX-RP separation. This is a measure of the complexity of the samples for MALDI-MS analysis.
2) The frequency of overlapping peptide ion clusters.
3) The total number of identified peptide species as a measure of the completeness of the peptide analysis and of ion suppression.

The results from these experiments are summarised in table 4.1.

The UV chromatograms from the 1D RP separation of the outer membrane protein digest shown in figure 4.6a reveals that the peptide elution profile is very

complex, with only partially resolved and overlapping peptide species. This is indicative that many peptides are co-eluting at the similar retention times. This was confirmed by mass spectrometric analysis, an example of which is shown in figure 4.7a. Similar analysis of other spectra revealed that between 10 and 30 peptides typically eluted in each fraction (table 4.1). Mass spectrometric analysis of all of the LC fractions revealed that there were 204 peptides detected in total in the fractions collected from the 1D RP separation. Mass spectrometric analysis of 20 selected fractions revealed the presence of 10 overlapping clusters. In comparison, the UV tracings from the 2D IEX-RP separation are less complex with better resolution of peptide species (figure 4.6b). This was confirmed by mass spectrometric analysis which showed that there were typically 5-15 peptides per fraction (figure 4.7b). Mass spectrometric analysis of all of the LC fractions identified 214 peptides in total (table 4.1). The 1M salt sample was not analysed due to loss of the sample during the experimental procedure. Importantly, only 1 overlapping peptide cluster was identified, so the frequency was much lower with 2D analysis.

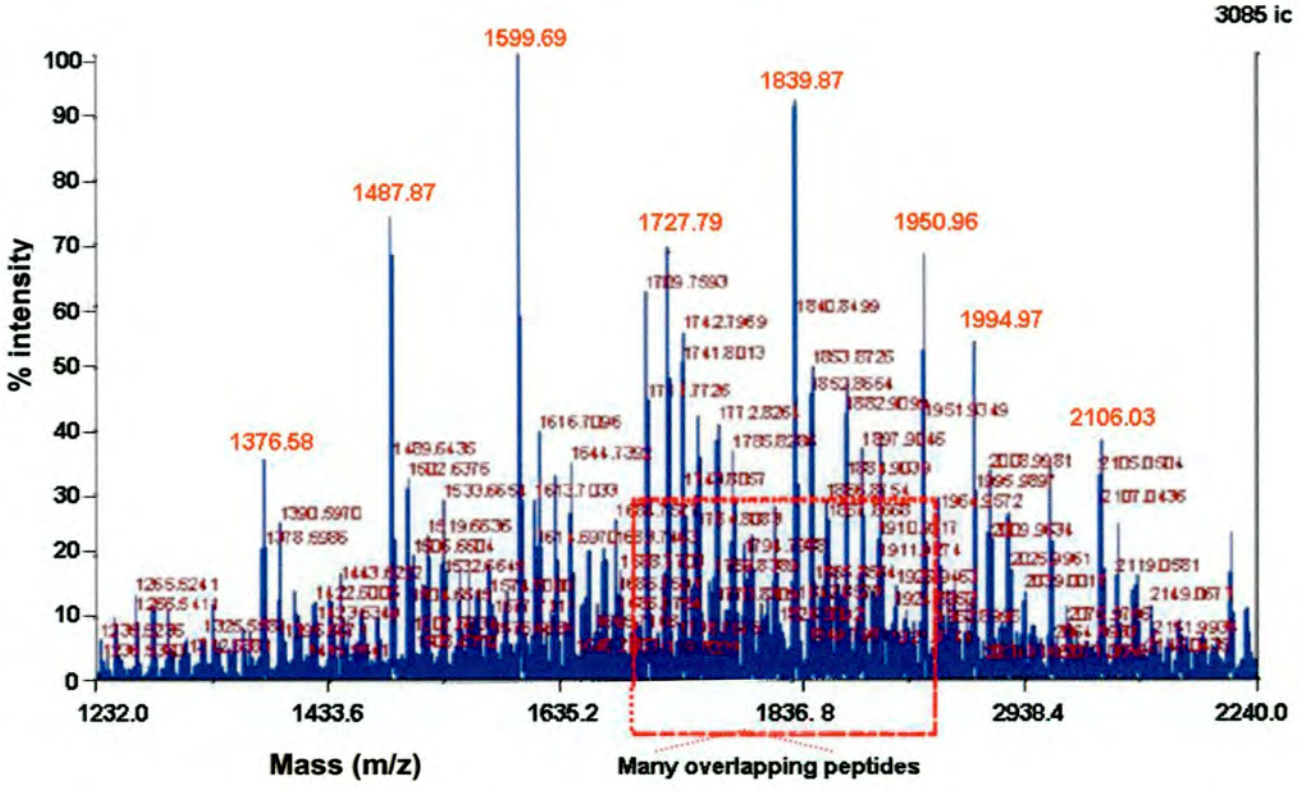


Figure 4.5 Mass spectrogram of a tryptic digest of *E. coli* outer membrane proteins

A mass spectrogram showing a Zip™ Tip of a tryptic digest of *E. coli* outer membrane proteins.

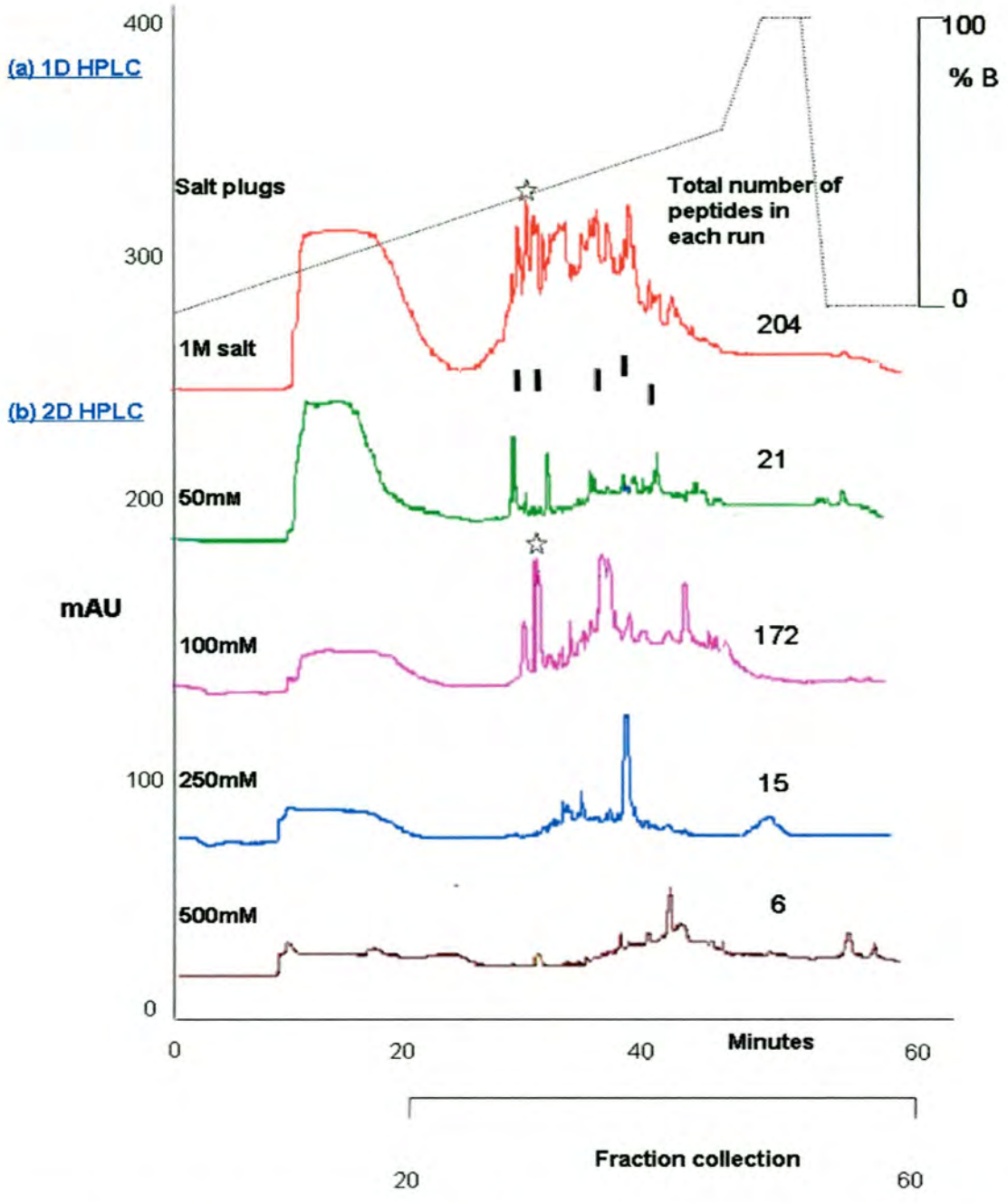


Figure 4.6 1D RP & 2D IEX-RP HPLC separation of a tryptic digest of outer membrane proteins

(a) UV tracing from the 1D separation of the outer membrane protein digest. **(b)** UV tracings obtained from the 2D separation of the outer membrane protein digest.

Legend

Star above fractions: HPLC fraction shown in figure 4.7

Dash below fractions: Fractions analysed in each run for overlapping peptides (numbers shown in table 4.2)

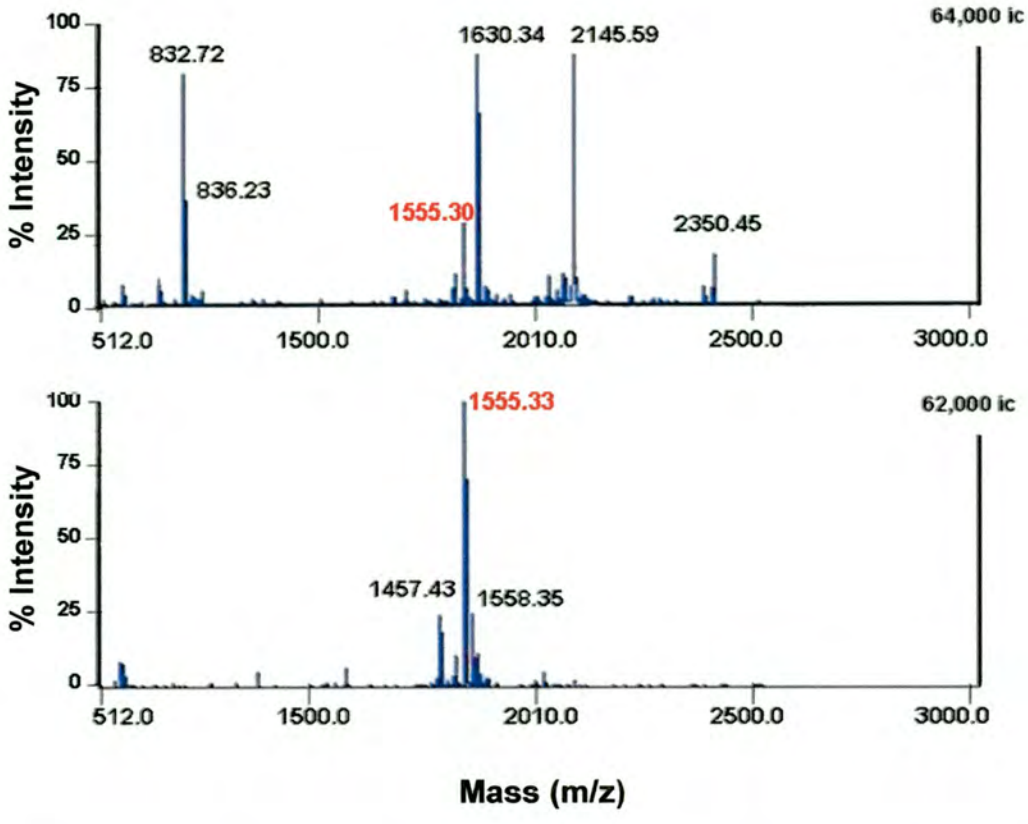


Figure 4.7 Mass spectrometric tracings from fraction number 30 following separation of a tryptic digest of *E. coli* outer membrane proteins

(a) Fraction 30 from the 1D RP separation of a tryptic digest of outer membrane proteins **(b)** Fraction 30 from the 100mM salt step in the 2D IEX-RP separation of a tryptic digest of outer membrane proteins.

Type of separation	Typical number of peptides co-eluting *	Number of overlapping peptide species **	Total number of peptides *
1D RP LC	10-30	10	204
2D IEX-RP LC	5-15	1	>214***

Table 4.1 Comparison of peptide separation and identification of labelled species following 1D and 2D HPLC separation

Legend

* Total number of peptides estimated by using the Voyager-DE PRO software to estimate the number of peptides in each HPLC fraction.

** Number of overlapping peptides counted in the HPLC fractions analysed for the 1D and 2D separation.

*** 1M salt plug not analysed due to loss of sample

4.3.3 Technical difficulties

Establishment of the nano-flow LC system and subsequent experiments carried out using the system was impeded by a number of technical difficulties that are summarised in table 4.2. These technical problems led to these experiments taking several months to complete. A recurrent problem was the presence of leaks and blockages that were hard to find in the system because of the low flow rates employed. There were also considerable difficulties in establishing correct synchronisation between the many different components of the system. In addition, there was a very steep learning curve on the way to being able to use the system efficiently.

Problem	Explanation	Outcome
System development	Learning curve to use software Overcoming communication failure	Takes time to be able to use LC system Failure of different components to start (e.g. pumps, fraction collector)
Poor performance of old equipment	(e.g.) Pump system	Pumps stop working during an LC run
Leaks & Blockages	Leaks often occur at fittings between capillary tubing Capillary system easily blocked by particles in sample Matrix crystal formation in capillary tubing	Loss of flow Sample redirected to waste Fraction collection onto MALDI-MS plate prevented
Inconsistent fraction size	Employment of a splitter system to achieve 200nl/min flow rate	Leads to peptide losses

Table 4.2 Problems encountered using the nano-flow LC system

4.4 Discussion

The experiments in this chapter were carried out to test whether the 2D nano-flow liquid chromatography system developed in the laboratory could improve the separation of complex peptide mixtures compared to conventional 1D chromatography sufficient to make carbon-13 labelling work for complex mixtures. At the initiation of this project, nano-flow LC was cutting edge technology, with limited published information on the best set-up to use in experiments. IEX followed by RP chromatography was chosen because they are largely orthogonal with respect to peptide properties, and well established chromatographic procedures for the separation of peptides. A number of other important issues had to be addressed during system development. These are discussed below.

A Magic Bullet strong cation exchange column was chosen for the first stage of separation. However, there are different types of C18 reverse phase column available to buy that differ in their dimensions, loading capacities and sensitivities. It was important that a column with a small internal diameter was used in order to maximise sensitivity of analysis. However, this was at the expense of reducing the amount of sample that could be introduced into the system, increasing analysis time and considerably increasing the difficulty in operation. The chosen compromise was a C18 column of 15cm in length and 75μm in diameter made by LC Packings. This column required a flow rate of about 200nl per minute, but at the initiation of this project, pump systems on the market could only pump at flow rates down to 50μl per minute. Therefore, it was necessary to install a flow splitter to obtain a flow rate of 200nl / min. The use of a splitter is not an ideal way to achieve low flow rates because the splitting ratio varies during the course of a run as the viscosity of the flow changes. This results in changes in the rate of flow that adversely influences fraction collection. Plus, slight blockages in the system result in the redirection of flow to waste as a consequence of an increase in backpressure. Difficulties in operating a splitter system were such that with the advent of pumps able to pump 200nl/min directly, the splitter has been a vital element in the system to be replaced, as happened soon after the end of my project.

The next important decision was to choose whether IEX and RP chromatography would be carried out sequentially 'online' or whether they would be

carried out as two separate experiments. This in turn would influence the decision of whether to use 'salt plugs' or a 'continuous salt gradient' to elute peptides from the ion exchange column prior to reverse phase separation. An online system is faster and requires samples to come into contact with fewer surfaces, potentially reducing analyte losses. Thus, the concern of peptide losses between 1D and 2D separation would be reduced. The disadvantage was the greater complexity of the instrumentation and control software, and the lower efficiency of ion exchange separation achieved by salt plugs compared to salt gradients (Nagele *et al.*, 2004). On line salt plugs were used in these experiments but in later experiments, the system was switched to using sequential analysis to improve the ion exchange separation achieved.

The results obtained in this study from the 1D RP and 2D IEX-RP LC separation of the outer membrane protein digest revealed that the 2D nano-flow LC system significantly improved separation of the complex peptide mixture and the subsequent identification of labelled peptides, without unacceptable peptide losses but at considerable cost in effort and time. There was also an increase in the total number of detectable peptides following 2D separation. This is likely to be the result of fewer peptide species co-eluting in a single fraction and consequent reduced potential for ion suppression effects. In fact, it was thought that the true number of peptides in the 2D run was probably greater than that observed. This was because the final 1M salt plug sample was unable to be analysed due to loss of the sample during purification.

For both 1D-RP and 2D IEX-RP separation, the UV tracings underestimated the total number of peptides detected by mass spectrometric analysis. For example, the UV tracing for the 1D separation of the tryptic digest shows that there are maybe between 20-30 peaks. However, mass spectrometric analysis revealed that there were in fact over 100 peptide species. This observation was also made for 2D separation, indicating that the addition of more salt steps may have better optimised 2D separation. Improving separation may potentially have the benefits of detecting a larger number of peptide species by overcoming ion suppressive effects and also preventing the occurrence of overlapping peptide clusters altogether. Indeed, 10 salt steps were used in the bulletin published by Michrom Biosciences, USA. This study

opted for fewer salt steps because of the many technical difficulties experienced at this time. In addition, although the employment of more salt steps leads to better separation, it also leads to diminishing returns in terms of the number of peptides detected, longer experiments and more spots to analyse.

The observation that a greater total number of peptide species was observed after 2D separation, suggested that any additional peptide losses was more than mitigated by improvements in mass spectrometry. There was, however, scope for improvement as it was observed that many peptides eluted over several salt steps, diluting some peptides into several fractions with a consequent decrease in sensitivity of analysis. Nagele *et al.*, 2004 has reported the same deficiency using the salt plug approach. Their solution was to design a liquid chromatography system that uses a semi-continuous salt gradient and two capillary entrapment columns, alternatively switched from the ion exchange effluent to the injection loop of a second dimension RP column. The time it might take to establish a two-enrichment column system was a concern, so in later experiments it was decided to use an off line continuous salt gradient system to further improve resolution.

Chapter 5 Identification of carbon-13 labelled tetanus toxin C fragment derived peptides following in-vitro lysosomal digestion

5.1 Introduction

The aim of the experiments in this chapter were to (i) validate the carbon-13 labelling technique in complex biological mixtures and to (ii) identify any peptide species generated from TTCF following digestion with lysosomes. The results may help in identifying enzymes apart from AEP that may be important in antigen digestion of TTCF and liberation of known T cell epitopes.

The results in the previous chapter revealed that the digest of *E. coli* outer membrane proteins was very complex, presumably consisting of many peptides derived from a wide range of proteins present within the outer membrane. Since lysosomal extracts are themselves complex mixtures, it was decided that it would be more favourable to initially examine the processing of a purified single protein such as TTCF rather than another complex mixture. The 47-KDa carboxy-terminal domain of tetanus toxin was chosen for several reasons. Firstly, it can be easily produced in large quantities as a correctly folded recombinant antigen. Secondly, there is a lot known about TTCF in respect to antigen processing and presentation. Studies have identified a number of T cell epitopes within tetanus toxin (Demotz, *et al.*, 1989; Ho, Mutch *et al.*, 1990; Reece, *et al.*, 1993; Raju, *et al.*, 1996; Diethelm-Okita, *et al.* 1997; Diethelm-Okita, *et al.*, 2000) and the initial processing steps of TTCF have already been elucidated (Manoury, *et al.*, 1998). These studies provide a groundwork of knowledge for this study to build upon before going on to study less well-characterised antigens.

Due to inadequacies in the 2D separation procedure carried out in the previous set of experiments, the system was changed for the experiments discussed in this chapter and is shown in figure 5.1. Peptide mixtures would be injected onto a cation exchange column and eluted into several fractions using a continuous salt gradient. The fractions would then be pooled according to complexity and separated by RP chromatography. In order to minimise potentially greater losses of an off line second dimension, thought was given to the handling of the first dimension fractions.

In particular, it was ensured that the fluid came into contact with the smallest possible surfaces.

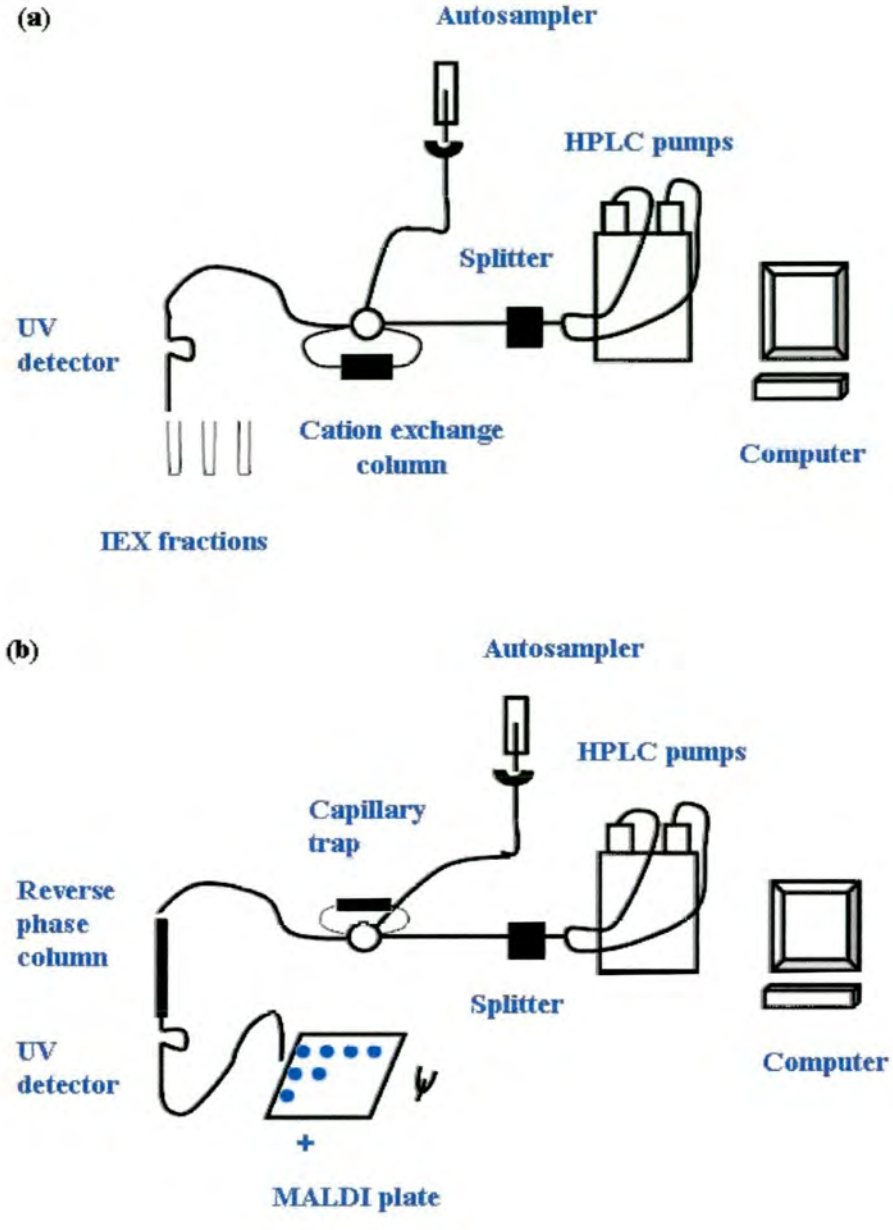


Figure 5.1 Set-up of the revised 2D HPLC system

(a) Diagram of the first dimension of separation, in which peptides are initially eluted from the IEX column using a continuous salt gradient and a number of fractions collected for reverse phase chromatography. **(b)** Diagram of the second dimension of separation, in which the sample is injected onto a capillary trap that is then subsequently switched in line with the RP column.

5.2 Objectives

- To produce carbon-13 labelled recombinant-TTCF (r-TTCF).
- To isolate functional lysosomes from EBV transformed Pala lymphoblastoid B cells.
- To confirm the major fragments of TTCF generated by lysosomal processing and to further identify any small (<3KDa) TTCF peptides generated by digestion.
- To thereby validate the carbon-13 labelling and purification strategies

5.3 Results

5.3.1 Production of Carbon-13 labelled TTCF

TTCF was expressed in carbon-13 supplemented minimal medium following induction with 0.2 mM IPTG. The results show that the expression of TTCF is under tight regulation, with TTCF not being detectable prior to the addition of IPTG (figure 5.2). This allowed practically all of the recombinant protein to be labelled by the addition of carbon-13 glucose at the time of induction. Tryptic fragments of TTCF enriched with 2.3% carbon-13 were distinguishable from their unlabelled counterparts throughout the m/z range 400-3000 Da (figure 5.3). A final yield of 5 mg / L was obtained following expression of the recombinant protein in M9 minimal medium. This yield was sufficient for the use of TTCF in the in-vitro digestion experiments.

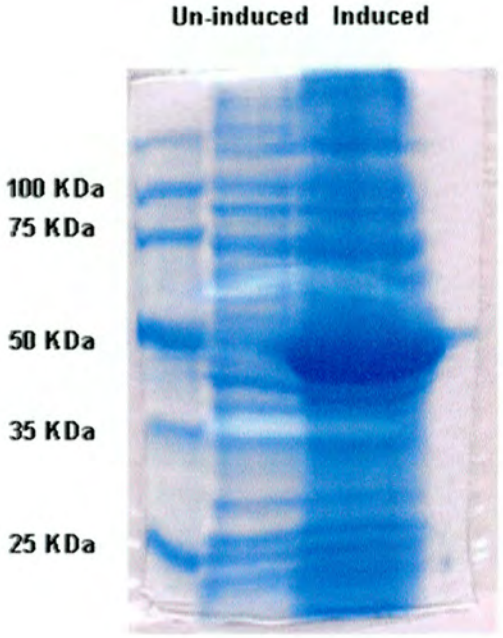


Figure 5.2 Expression of TTCF in M9 minimal medium

The TCP from the induced and uninduced bacterial cultures was run out on a 12% SDS-PAGE gel. *E. coli* expressing TTCF was grown to an OD_{550} value of 0.5-0.6 and TTCF production induced by the addition of 0.2 mM IPTG. The culture was harvested after 3 h and 10 µl of the total cell protein from both the uninduced and induced bacterial cultures analysed by SDS-PAGE analysis. The expected Mr of TTCF is 47 KDa.

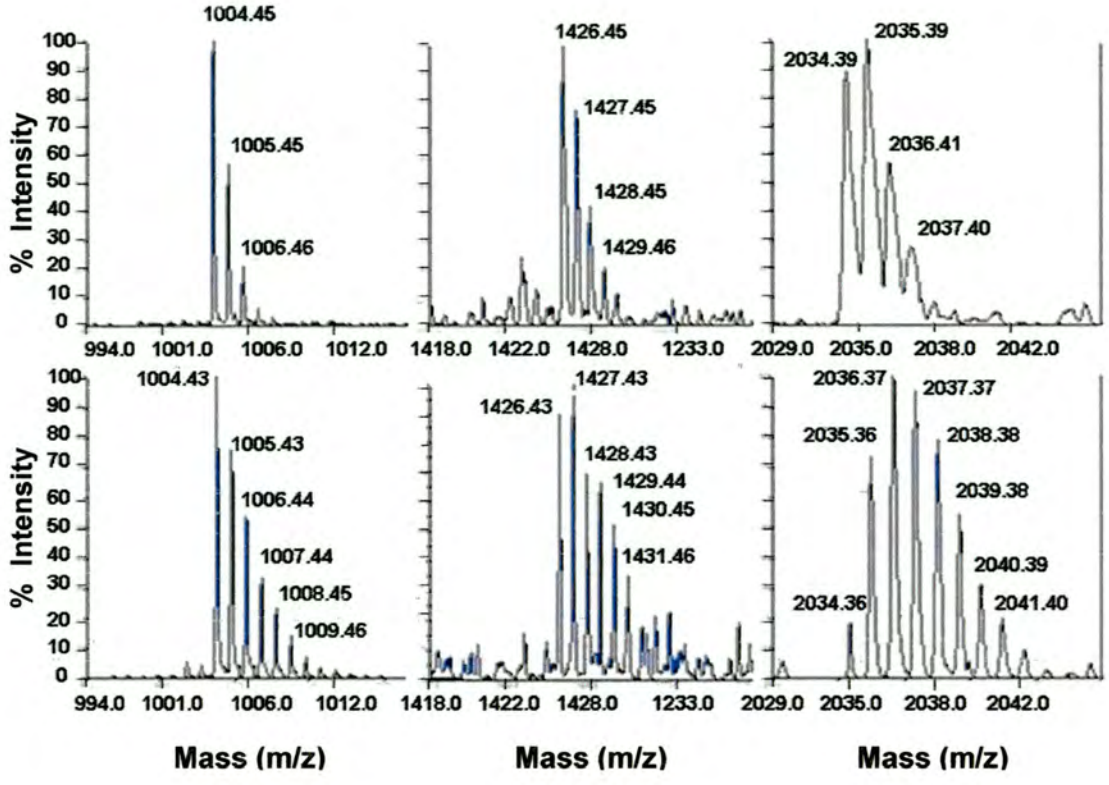


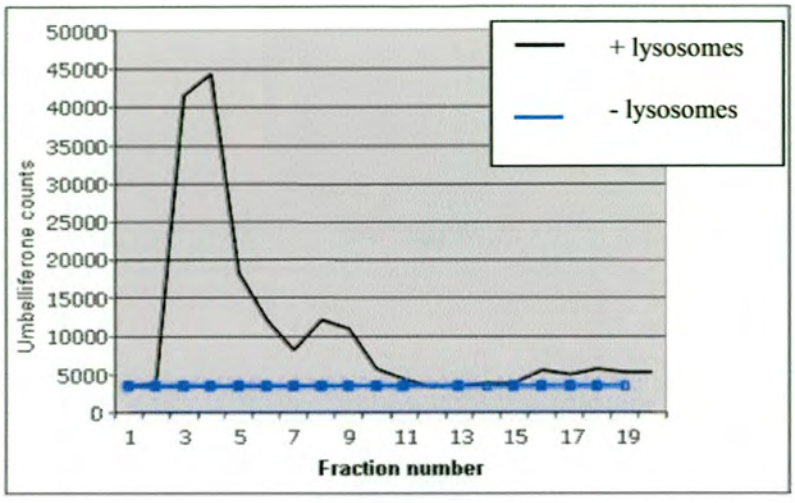
Figure 5.3 Carbon 13 enrichment of TTCF

The isotopic pattern for 3 TTCF tryptic peptides of differing m/z value is shown for a) The un-enriched peptide species containing 1.1% carbon-13 that is the natural abundance of carbon-13 found in nature and **b)** Following enrichment with 2.5% carbon-13

5.3.2 Production of lysosomes

Lysosomes were purified from human lymphoblastoid B cells and separated by density gradient centrifugation. The fractions containing lysosomes were determined using a fluorescent enzyme assay detecting the activity of N-acetyl-β-D-glucosaminidase (figure 5.4). This enzyme acts upon the substrate 4-methylumbelliferone to release the fluorescent compound, umbelliferone and was chosen because N-acetyl-β-D-glucosaminidase is active only in terminal lysosomes (Casciola-Rosen and Hubbard 1991). Results from the assay revealed that the highest umbelliferone counts were consistently given by fractions 3 to 5, with lower counts being observed in fractions 6 to 10. In all cases, the first 10 fractions were pooled together and subsequently used in digestion assays. Using a protein assay, it was estimated that 10^8 cells gave a yield of 100μg of lysosomal protein following the extraction procedure.

a)



b)

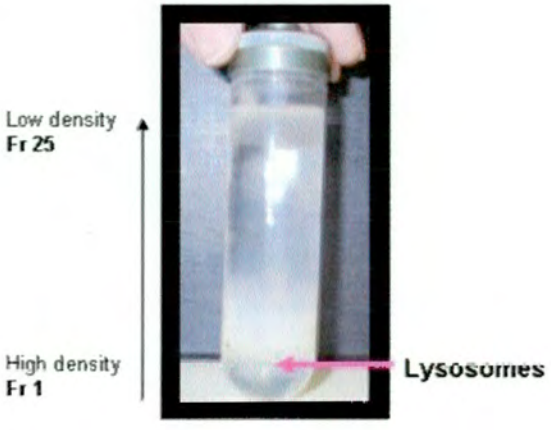


Figure 5.4 Determination of lysosomal fractions following sub-cellular fractionation of B cells on a 27% percoll gradient

Three millilitres of 10^6 / ml of broken B cells were applied to a 27% percoll gradient and centrifuged for 1h at 23,000 rpm to separate sub-cellular organelles based upon their density **a)** Graph to show which fractions contained lysosomes by measurement of N-acetyl-β-D–glucosaminidase **b)** Photo showing the sub-cellular components of B cells following percoll separation. Lysosomes, as determined in the enzyme assay, are denoted.

5.3.3 Lysosomal digestion of TTCF

TTCF was digested in-vitro using increasing amounts of lysosomal lysate (Figure 5.5). The results revealed that lysosomal digestion of TTCF yielded five large fragments distinguished in a Coomassie-stained 16% tris-tricine gel. These larger fragments persisted following digestion with the highest amounts of lysosomes. To study the lower molecular weight digestion species (<3KDa) not detectable using SDS-PAGE, a large scale two stage experiment was set up in which 50 μg of TTCF was initially digested at pH 5.5 for 4 h and then subsequently at the lower pH 4.5 for a further 4 h (figure 5.6a). The decrease in pH was carried out in an attempt to reflect acidification observed in the endosomal pathway. Lower molecular weight digestion products were extracted by ultrafiltration over a 3KDa cut off membrane and examined by MALDI-TOF. They comprised of an extremely complex mixture of peptides with m/z values between 500-1500 Da (figure 5.7a). This corresponds to peptides between 5 to 15 amino acids in length. Therefore, as well as the large fragments seen on SDS-PAGE, the lysosome – TTCF digests contained a highly complex mixture of low molecular weight peptides.

To investigate whether any of the low molecular peptides were TTCF derived, the digestion extracts were separated using 2D IEX-RP chromatography and analysed by MALDI-TOF. Disappointedly, although there were lots of different peptides, MALDI-TOF signals were weak. This result was surprising as the amount of TTCF used in the experiment was such that TTCF peptides should be detectable. One reason for non-detection could be the high stability of the larger fragments and consequent slow production of smaller fragments that could be detected by MS analysis. Another could be rapid destruction of any released peptides. The experiment was therefore repeated using more antigen (500 μg) and a shorter digestion time (figure 5.6b). Analysis of the unseparated digestion products by MALDI-MS revealed abundant quantities of hundreds of peptides with slightly higher m/z values than the preceding, longer duration digestion experiment. Most peptides had m/z values in the range 1000-2500 Da (figure 5.7b), corresponding to peptides of 10 to 20 amino acids in length. Peptides in this range are substrates for MHC class II molecules. These peptides were examined in detail to identify any that derived from TTCF.

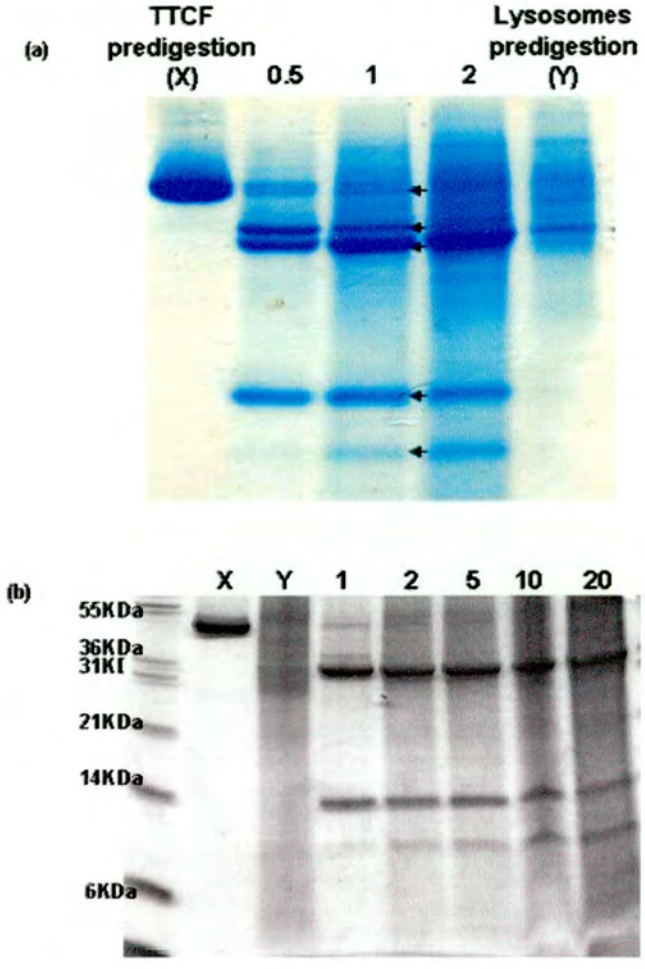


Figure 5.5 Digestion of TTCF with increasing amounts of lysosomes

TTCF digestion products were separated using a 16% Tris-Tricine gel. **a)** 10μg of TTCF was digested with 0.5– 2μg of lysosomes purified from EBV-transformed lymphoblastoid B cells (PALA B cells). Digestion products are indicated by black arrows **b)** 10μg of TTCF was digested with 1–20μg of lysosomes.

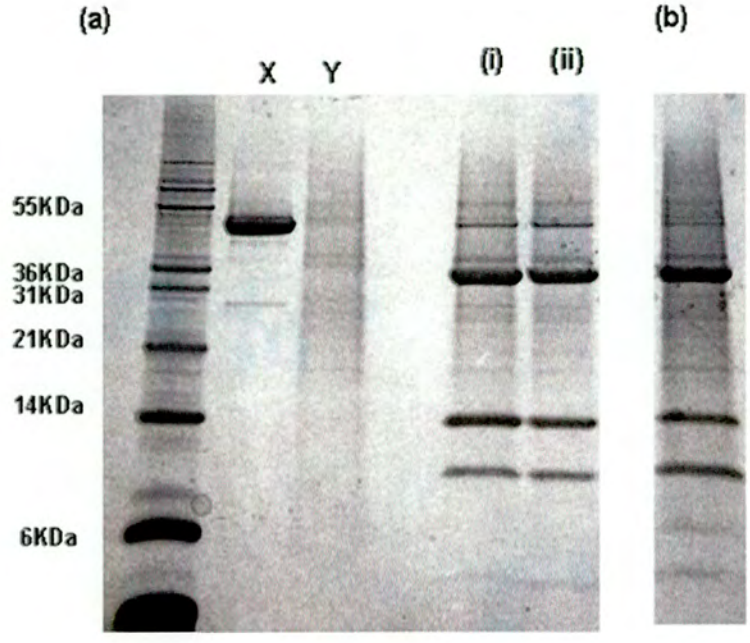


Figure 5.6 Digestion of TTCF using different digestion parameters
In all digestion experiments described below, the equivalent of 10µg is analysed using a 16% Tris-Tricine gel. **a)** In the first experiment, 50µg of TTCF was digested using 20 µg of lysosomes for **(i)** 4h at pH 5.5 and then for a further **(ii)** 4h at pH4.5 with the addition of 2µg of fresh lysosomes. **b)** In a second experiment, 500µg of TTCF was digested with 100µg of lysosomes for 4h at pH4.5

Legend
X: TTCF predigestion
Y: Lysosomes predigestion

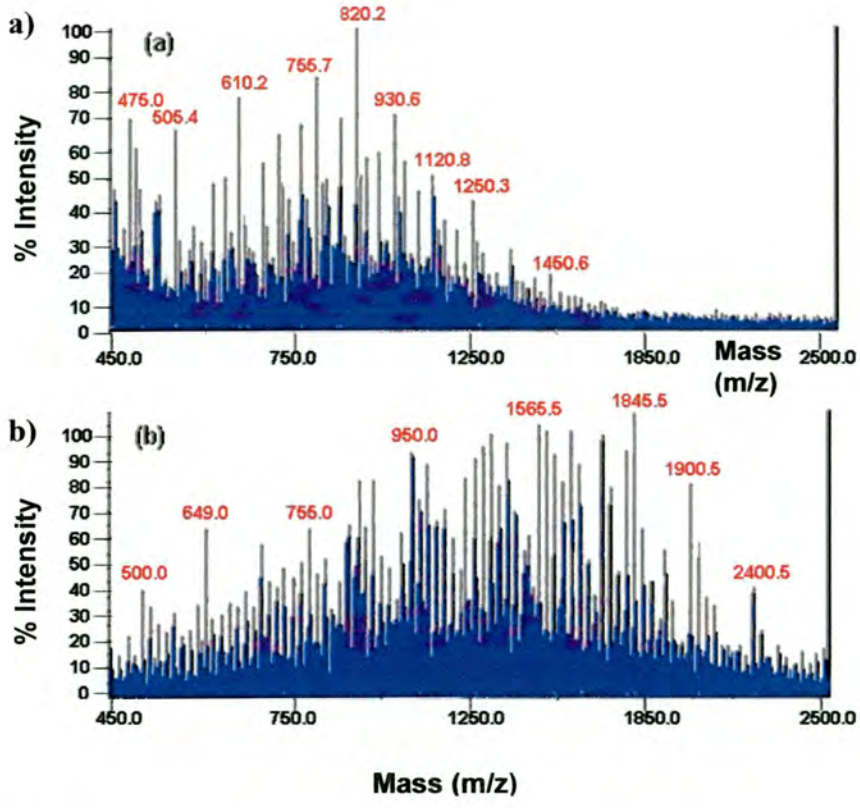


Figure 5.7 MALDI-MS analysis of unseparated digestion products following lysosomal digestion of TTCF

Mass spectrograms of peptides in the m/z range 400-2500 Da are shown for the unseparated digestion products following **a)** the 8h digestion period shown in 5.5a (ii) and **b)** the 4h digestion period shown in 5.5b. For both samples, the digestion mixture was passed initially through a 3KDa cut-off filter (Amicon). A small proportion was subsequently concentrated and desalted using a C18 Zip™ Tip, and analysed by MALDI-TOF MS to give the mass spectrograms shown above. The MS conditions under which spectra were collected and processed were the same for both samples.

5.3.4 Analysis of peptides <3000 Da recovered from lysosomal digests of TTCF

The peptide mixture was highly complex so it was decided that 2D IEX-RP separation would be employed to generate a large number of less complex fractions for analysis. UV tracings from the 2D IEX-RP separation contained large numbers of partially resolved peaks (figure 5.8). However, separation was fairly successful in that mass spectrometric analysis of the fractions revealed between 20 and 100 peptides at which complexity, ion suppression and overlapping clusters should be less frequent.

After extensive examination of the spectra, two peptides were identified as having a carbon-13 'labelled' isotopic appearance (figure 5.9 & table 5.1). Several other isotope clusters could have been labelled or accounted for as overlaps. The m/z of the putative TTCF derived peptides each matched the predicted m/z of several TTCF subsequences. More precise measurement of m/z values using internal standards is often sufficient to distinguish the correct sub-sequences (Zou *et al.*, 2003). However, this approach was unsuccessful here because the putative TTCF derived peptide ions were of too low abundance to be seen in spectra that included internal peptide standards. The best that could be achieved was a close external calibration which gives rise to much larger errors between the experimental and predicted peptide m/z values.

The peptide of mass 2208Da matched to three TTCF sequences whilst the peptide of mass 2060.1 matched to two TTCF derived sequences. All predicted matches for both of these peptides matched to peptides of 17 to 19 amino acids in length, which is within the characteristic size for peptides eluted from MHC class II molecules. These matches are shown superimposed upon the TTCF sequence in figure 5.10. None of these candidates matched to previously elucidated cut sites for enzymes implicated in antigen processing of TTCF or to known T cell epitopes (Hewitt *et al.*, 1997; Manoury *et al.*, 1998). Overall, these results indicated that lower molecular weight TTCF fragments can be identified from their labelled appearance but are surprisingly scarce in the lysosomal digest, despite the vast relative abundance of TTCF in the digestion mixture which was present at least 10 times the amount of any other protein.

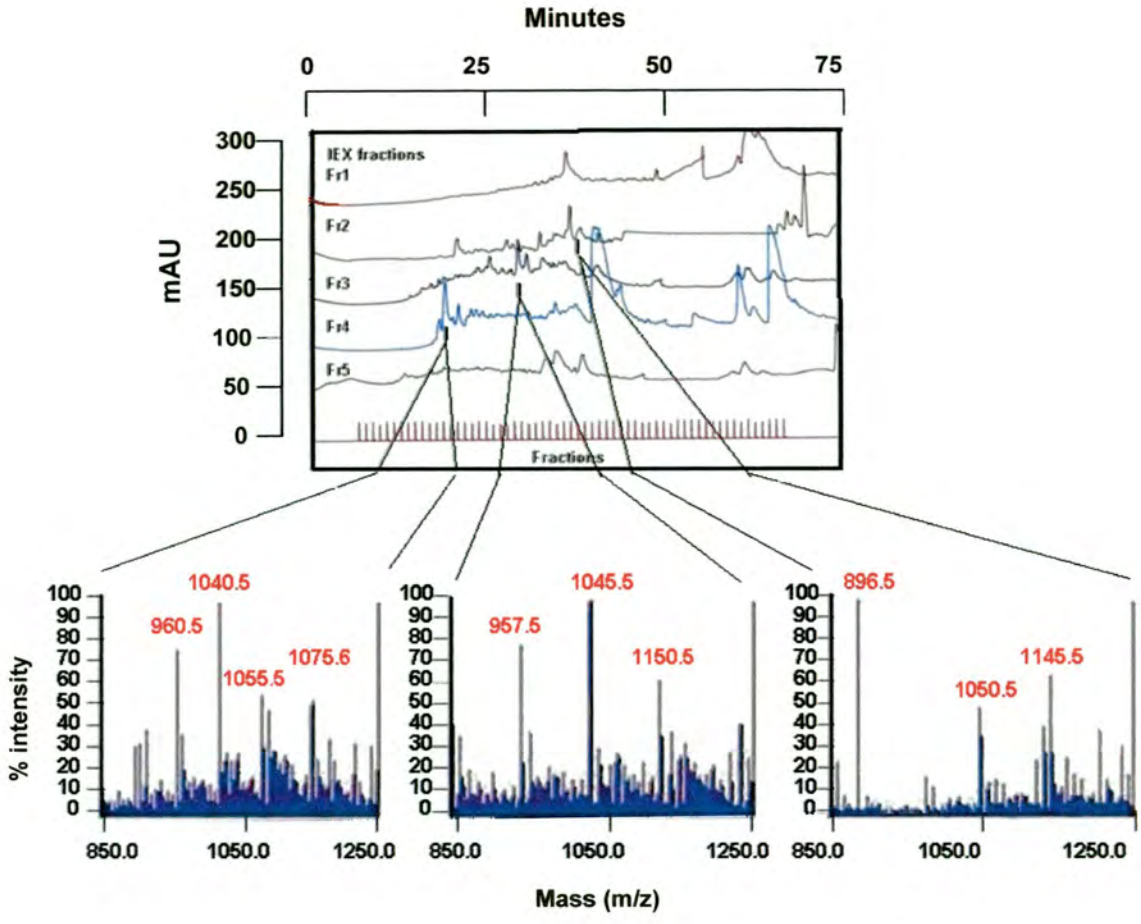


Figure 5.8 2D IEX-RP nano-flow LC separation of digestion products following lysosomal digestion of TTCF

a) UV tracings from the RP separation of the 5 IEX fractions obtained from IEX separation of the lysosomal digestion products. **b)** Mass spectrograms of 3 HPLC fractions following RP separation showing peptides with m/z values between 850 – 1250 Da.

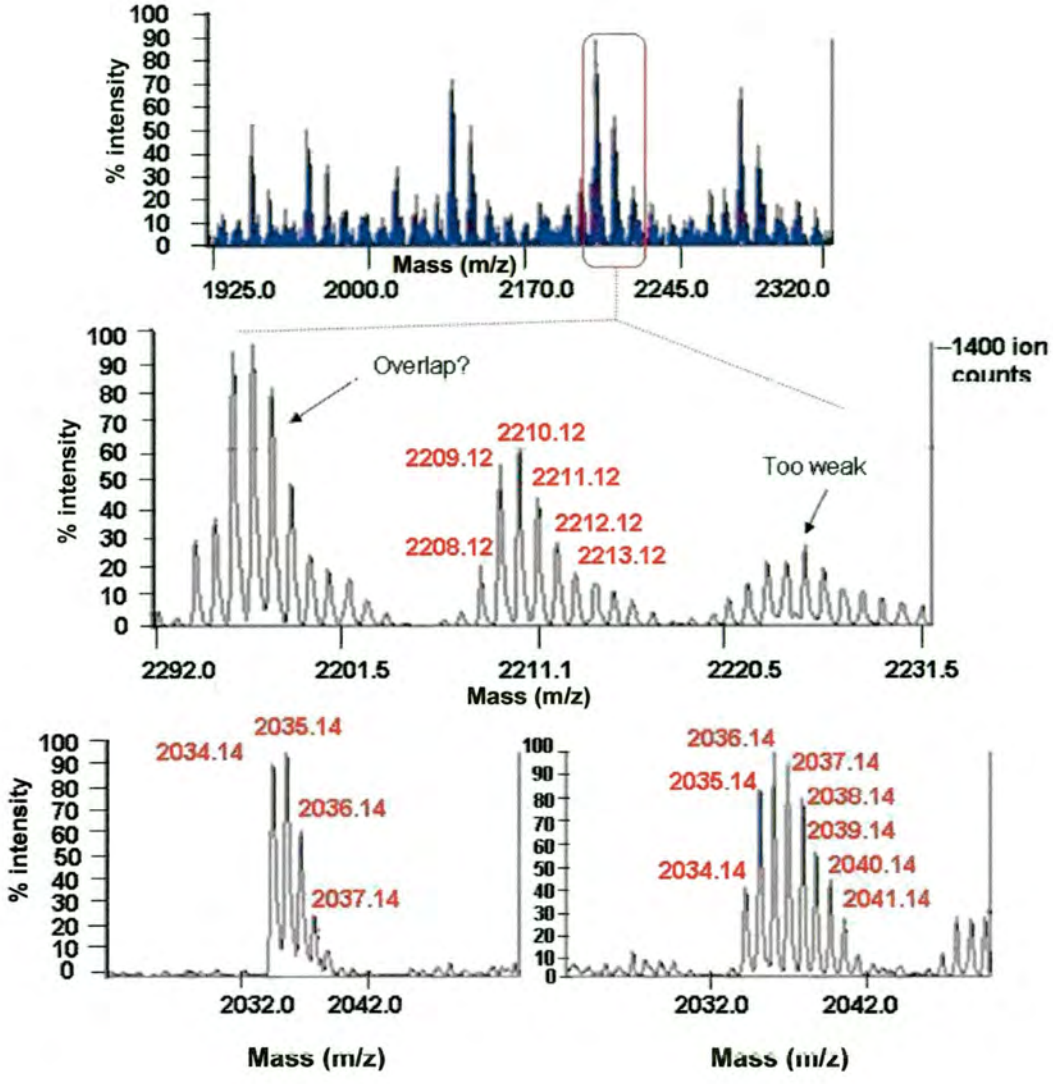


Figure 5.9 Identification of a peptide with a carbon-13 'labelled' isotopic appearance following 2D IEX-RP nano-flow LC separation of the digestion products obtained from lysosomal digestion of TTCF

a) MALDI-TOF MS analysis of fraction 30 from the RP separation of IEX Fr 4 **b)** Mass spectrogram of fraction 30 showing a peptide with a m/z value of 2008.1 which was identified as having a 'labelled' isotopic appearance. There are two other potentially labelled peptides in the spectra, one of which is most probably accounted for by an overlapping peptide cluster and the other of which is too weak to determine whether it is a labelled peptide. **c)** Mass spectrogram of a peptide of similar m/z value of 2034.1, generated by the trypsin digestion of unlabelled and labelled TTCF.

Labelled peptide (m/z)	TTCF-subsequences with matching calculated m/z (+/- 100 ppm)	
2208.3213	2208.1044 2208.1693 2208.1560	TLK<DSAGEVRQITFRDLPDKFN>AYL WVF<ITITNDRLSSANLYINGVLM>GSA NAP<SYTNGKLNIYYRRLYNGL>KFI
2060.1816	2060.0771 2060.0382	IWT<LKDSAGEVRQITFRDLPD>KFN RCN<NNNQYVSIDKFRIFCKA>LNP

Table 5.1 TTCF derived sequences that could account for the 2 identified labelled peptides

The masses of the 2 peptides with a labelled isotopic appearance were searched against a list of calculated prelisted m/z values for all possible subsequences of TTCF. The maximum error allowed between experimental and predicted masses was 100ppm. The start and end of the predicted match is denoted by the symbol '<'.

Histidine tag

MGHGHHHHHHHHHSSGHIEGRHMLDNEEDIDVILKKSTILNLDINNDII

SDSGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMF

NNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKG

NNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLY

INGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALN

PKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITD

YMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIK

LYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRD

LKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK

ILGCDWYFVPTDEGWTND

Figure 5.10 TTCF sequence matches for the two peptides identified as having a carbon-13 labelled isotopic appearance

The TTCF sequences which matched by mass (+/- 100ppm) to the peptides identified as having a labelled isotopic appearance are shown by lines above (for 2060) or below (for 2208) the sequence of TTCF denoted in single letter amino acid code. A histidine tag MGHGHHHHHHHHHSSGHIEGRHI precedes the TTCF sequence.

5.3.5 Digestion of TTCF with different subcellular fractions

One concern was that the scarcity of lower molecular weight digestion products might be a consequence of unintended selection of subpopulations of processing compartments that did not contain all of the enzymes required to digest the larger TTCF fragments into lower digestion products. Alternatively, the selected population may have contained too many enzymes. It is known that early endosomes mature into late endosomes and eventually fuse with lysosomes. During this maturation, there is a gradual decrease in pH from 6.5 to 4 with an accompanying increase in enzyme activity.

To examine the TTCF processing capacity of endo / lysosomal organelles of different density, TTCF digestions were set up using different sets of fractions from the separation of the subcellular components of B cells (figure 5.11). The results revealed that digestion of the parent TTCF molecule is seen with the pooled fractions 1-7, 8-10, and 11-13. Conversely, no digestion was observed using fractions 14-25. In digestions using fractions 1-13, the same large digestion fragments were detectable using a 16% tris-tricine gel. However, it appeared that digestion became progressively weaker going from fractions 1-7 to 11-13 and importantly no additional digestion products were detected during digestion by the different processing compartments.

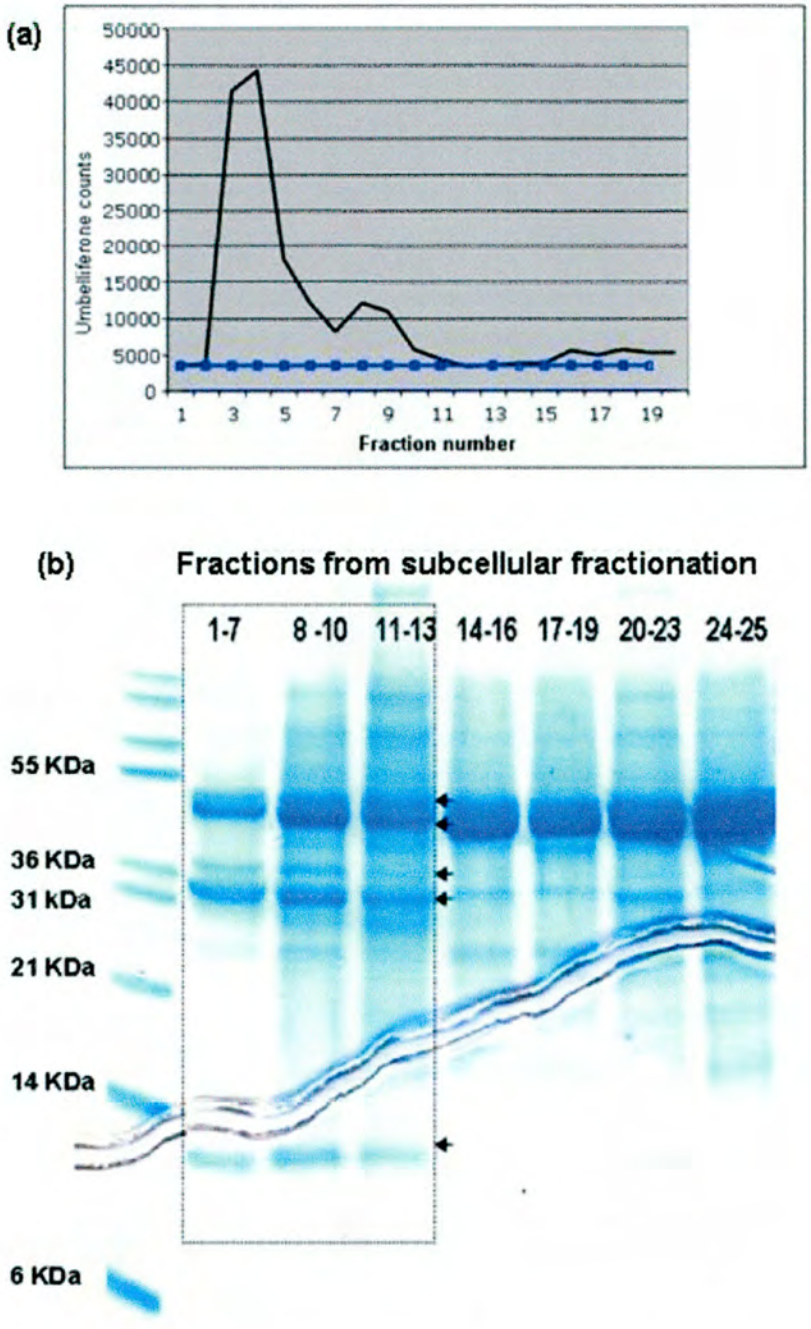


Figure 5.11 Digestion of TTCF using different fractions obtained by the percoll based fractionation of sub-cellular organelles from broken B cells

(a) Shows umbelliferone counts for subcellular fractions of B cells obtained by percoll based separation of homogenised B cells. **(b)** Fractions were pooled together according to their N-acetyl-β-D–glucosaminidase activity (shown in a) and 10μg of TTCF digested using different 'pooled' fractions. The digestion products were separated using a 16% Tris-Tricine gel. Arrows denote digestion products.

5.4 Discussion

In the experiments described in this chapter, the aim was to identify carbon-13 labelled TTCF digestion products following in-vitro digestion with lysosomes isolated from EBV transformed B cells. It was demonstrated that carbon-13 labelled TTCF could be successfully expressed in M9 minimal medium at quantities sufficient for use in such large-scale experiments. Lysosomal preparations were also successfully purified by a combination of cell homogenisation followed by Percoll based density centrifugation to separate the subcellular components (Davidson *et al.*, 1990). This method was chosen because it was previously used by Manoury *et al.*, (1998) to isolate lysosomes for in-vitro studies in studying the initial digestion steps of TTCF.

The problem of identifying low molecular weight (800-4000 Da) TTCF digestion products was tackled by using 2D nano-flow liquid chromatography to separate out the complex digestion mixture prior to characterisation of the processing products by MALDI-MS. In the large-scale digestion experiment in which 500 μg of TTCF was digested by 100μg lysosomes, the peptides ranged from between 10 to 20 amino acids in length. These are the sizes of peptides eluted from MHC class II molecules and encouraged the search for TTCF fragments. However, only two peptides were finally identified as having a labelled isotopic appearance. It was hypothesised that few TTCF peptides may have been identified because not all of the processing compartments required for complete digestion were used in the in-vitro experiments. However, digestion using other subcellular compartments did not lead to more complete digestion. The observation that digestion became progressively less extensive going from fractions 1-13 is consistent with the idea that harsher digestion occurs in lysosomes and late endosomes (most dense fractions) compared to earlier endosomes (less dense fractions).

It should be noted that both of the peptides identified as being labelled peptides were present in complex fractions containing up to 100 peptides so it is possible that they were part of overlapping peptide clusters that resembled a labelled peptide. In addition, both peptides gave rise to weak signals of below one thousand ion counts during mass spectrometric analysis, which decreases the certainty of distinguishing labelled peptides. The weak signals obtained for each of the peptides

meant that the peptides could not be further characterised by more precise m/z measurement of sequencing. Confirmation of the peptides would require the experiment to be repeated on a larger scale and sequence analysis by tandem mass spectrometry.

Possibly, an alternative lysosomal isolation technique may have improved the experiment by reducing the complexity of the lysosomal autolysis mixture. In the method used in this study, the detergent CHAPS was used to solubilise the lysosomal membrane prior to digestion. Thus, proteins from the lysosomal membrane may have been solubilised and contributed to background peptides. An alternative method to isolate pure lysosomal enzymes uses a combination of differential centrifugation and hypotonic lysis of lysosomes (Schroter *et al.*, 1999). In this method, the lysosomes are burst open by the addition of water and subsequently removed by centrifugation. Pure lysosomal enzymes are then recovered by a second centrifugation step. Using this technique, which would lead to improved separation and less complex fractions, may significantly reduce background mess.

The principal conclusion that TTCF low molecular weight peptides are scarce was challengeable because the sensitivity of the carbon-13 labelling technique in biological extracts has not been determined. Previous to this project, the carbon-13 labelling technique had been shown to be highly sensitive and specific in controlled mixtures of labelled and unlabelled peptides (Zou *et al.*, 2004). In these mixtures, the relative abundances of labelled peptides were higher and the peptide mixtures less complex than those examined here. In addition, the high laser intensity required to detect the peptides in many fractions led to a noisier baseline than that usually observed with more concentrated samples. It may therefore be the case that the carbon-13 labelling technique can not be used to identify low level peptide species in peptide mixtures of such complexity The experiments in the following chapter were set up to test this hypothesis.

Chapter 6 Identification of tetanus toxin C fragment-derived carbon-13 labelled peptides in complex peptide mixtures

6.1 Introduction

In the experiments in this chapter, the sensitivity of the carbon-13 labelling technique in identifying carbon-13 labelled TTCF peptides in complex mixtures was tested. It was decided to test sensitivity independent of digestion by 'spiking' pre-digested TTCF peptides firstly into a peptide mixture and then into lysosomal digests. The experiments were prompted by data from the previous chapter in which TTCF derived peptides were unexpectedly scarce when assessed by 2D IEX-RP nano-flow HPLC and MALDI-TOF. The decision was made to separate the mixtures in the proceeding set of experiments using 1D chromatography in preference to 2D analysis in order to reduce the amount of data to analyse and thus the time scale of the experiment.

6.2 Objectives

- To characterise carbon-13 labelled TTCF peptides generated by trypsin digestion.
- To examine the effectiveness of carbon-13 labelling at identifying TTCF peptides spiked at low relative abundance into complex peptide mixtures generated by trypsin digestion of 3 proteins.
- To compare the effectiveness of carbon-13 labelling at identifying TTCF peptides spiked at low relative abundance into lysosomal extracts that have been permitted to autolyse for 4 h at 37 °C and contain a complex mixture of peptides and other lysosomal constituents.

6.3 Results

6.3.1 Characterisation of tryptic fragments of labelled TTCF derived peptides

The m/z values of peptides generated by tryptic digest of TTCF were measured after internal calibration against 2 autolysis peptides of trypsin. Eight of these peptides were identified as having a labelled isotopic appearance and matched to a TTCF sequence following calibration (table 6.1). It was found that the peptides

gave rise to variable signal intensities during mass spectrometric analysis, ranging from 1000 to 10,000 ion counts. Sensitivity and specificity data showed that the carbon-13 test was 100% sensitive; identifying all TTCF derived peptides as being labelled (table 6.2). The test was also shown to be highly specific, with only 1 non-TTCF derived peptide being identified as having a labelled appearance. This peptide gave rise to a very weak ion count, which may have accounted for the error as weak peptides often resemble labelled peptides. Five distinct ions in the mass spectrum did not match to the TTCF sequence nor did they have a labelled isotopic appearance. These unlabelled non-TTCF derived peptides were assumed to be from contaminants that are often present following trypsin digestion of gel slabs.

Mass	Ion counts (ic)	Sequence match	Error (ppm) +/- 100ppm
664.3692	6000	EVR<QITFR>DLP	12
721.3956	7500	LYK<KMEAVK>LRD	5
841.4511	>10000	NGK<LNIYYR>RLY	6
1004.4970	>10000	FLR<DFWGNPLR>YDT	2
1135.5	3000	YPK<DGNAFNNLDR>ILR	1
1291.6960	700	ILR<VGYNAPGIPLYK>KME	13
1426.5169	900	IKR<YTPNNEIDSFVK>SGD	5
2034.0795	900	DDK<NASLGLVGTHNGQIGNDPNR>DIL	33

Table 6.1 Characterisation of labelled TTCF peptides following trypsin digestion
The peptide masses generated by trypsin digestion of TTCF were searched against a list of calculated prelisted m/z values for all possible subsequences of TTCF. An error of +/- 100 ppm was tolerated between experimental and predicted values.

	TTCF derived	Non-TTCF derived
Labelled appearance	**8 (a)**	**1 (b)**
Unlabelled appearance	**0 (c)**	**5 (d)**

Sensitivity = a/ (a+c) = 8/8 =100%

Specificity = d/ (b+d) = 5/6 = 83%

Table 6.2 Data showing sensitivity and specificity values for the carbon -13 labelling technique in the identification of labelled TTCF peptides following tryptic digestion of TTCF

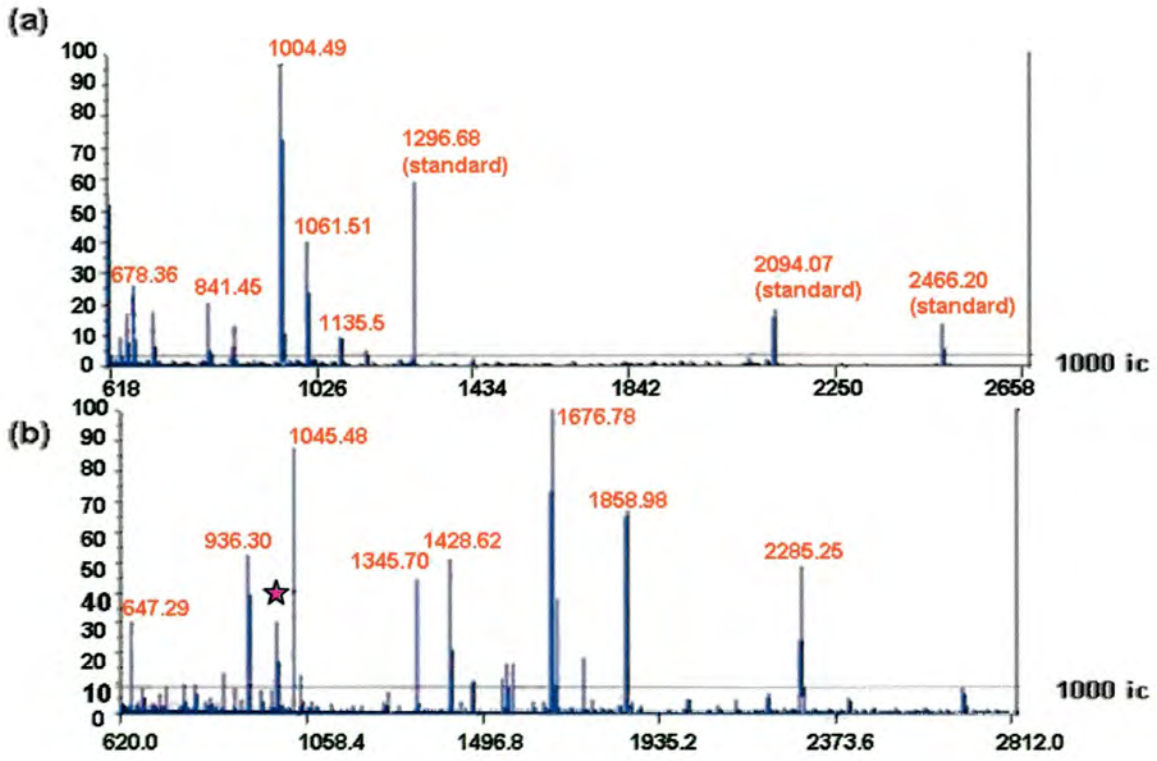
Sensitivity = The proportion of TTCF derived peptides which are detected as having a labelled isotopic appearance

Specificity = The proportion of non – TTCF derived peptides that are detected as having an unlabelled TTCF appearance.

6.3.2 Identification of tryptic fragments of labelled TTCF in mixtures with more abundant unlabelled peptides

To test the capacity of the technique to distinguish labelled from unlabelled peptides, 10 picomoles of tryptic fragments from TTCF were mixed with 50 picomoles of a mixture of peptides generated by trypsin digestion of three unlabelled proteins (figure 6.1). Mass spectrometric analysis of the unseparated mixture identified only 1 out of the 8 TTCF fragments. This peptide had a m/z value of 1004.5 and had previously given rise to the highest ion counts in the tryptic digest of TTCF, prior to mixing with the unlabelled peptides shown in table 6.1.

MALDI-TOF analysis of fractions following RP separation of the mixture permitted the detection of all 8 labelled TTCF peptides (figure 6.2). The benefits of RP separation were mostly a consequence of improved ionisation of the TTCF peptides in less complex mixtures, with between 2-5 peptides being observed in each fraction. In every case, TTCF derived peptide ions were identified with a labelled isotope pattern in fractions collected at retention times identified previously to those fractions containing the respective peptides in RP-separation of the TTCF mixture alone.



☆ **1004.49 (labelled TTCF peptide)**

Figure 6.1 Identification of carbon-13 labelled characterised TTCF peptides in a complex mixture of labelled and unlabelled tryptic peptides

a) Mass spectrogram of 10 pmoles of a tryptic digest of carbon-13 labelled TTCF **b)** Mass spectrogram of 10 pmoles of the same TTCF tryptic digest mixed with 50 pmoles of a tryptic digest of three unlabelled proteins (BSA, Hb, chicken egg lysozyme).

Star denotes a labelled TTCF peptide of mass 1004.49 which is still detectable following mixture of labelled and unlabelled peptides

ic: ion counts

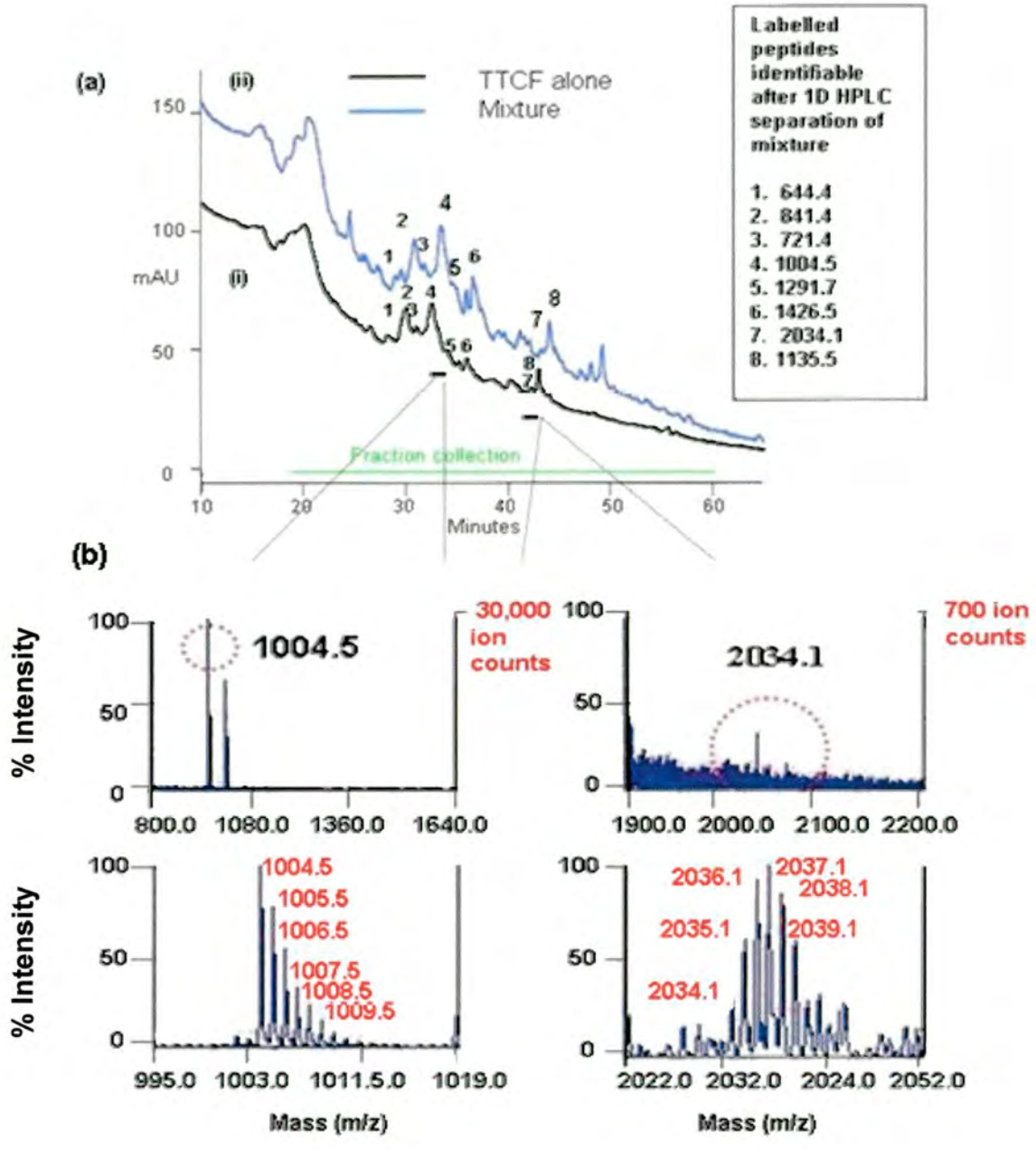


Figure 6.2 Identification of carbon-13 labelled TTCF peptides in a complex mixture of labelled and unlabelled tryptic peptides following RP separation

a) UV tracings from the RP nano-flow LC separation of 10 pmoles of the labelled TTCF tryptic peptides **(i)** alone **(ii)** or mixed with 50 pmoles of a tryptic digest of three unlabelled proteins. Numbers above the tracings indicate the presence of a labelled peptide within the fraction collected at that retention time. The identity of the labelled peptide is shown in the table next to the UV tracings. **b)** Mass spectrograms for two of the fractions (Fr 23 & Fr 45) from the RP separation of the mixture of labelled and unlabelled peptides containing the labelled TTCF peptides 1004.5 and 2034.1 respectively. For both peptides, the bottom spectrogram in each panel is a blow up of the peptide shown in the spectrogram above, showing its labelled isotopic pattern.

6.3.3 Identification of tryptic fragments of labelled TTCF in lysosome autolysis mixtures

To test the capacity of the technique to identify labelled peptides in lysosomal autolysis digests, 10 micrograms of lysosomes were incubated at 37°C for 4 hours, and then spiked with 10 pmoles of TTCF and immediately injected into the HPLC system. MALDI-TOF analysis of an aliquot taken before RP separation contained a complex mixture in which 0 of the 8 labelled TTCF fragments were detected. The mixture was extremely complex, containing hundreds and possibly thousands of peptides (figure 6.3). RP separation of the mixture gave a complex chromatogram (figure 6.4a) in which several peaks had retention times identical to 5 of the 8 TTCF fragments. Mass spectrometric analysis identified 5 out of the 8 labelled peptides in the fractions collected at the expected retention times, as shown for the peptide of m/z value 2034.1 in figure 6.4bi. One of these peptides, 1135.5, was found to be part of an overlapping cluster, preventing the recognition of the labelled isotopic pattern (6.4bii).

It was thought that ion suppression from the sheer number of peptides in many of the fractions may have been responsible for 3 / 8 of the labelled peptides not being detectable following RP separation of the mixture. Indeed, all 3 of these peptides had given rise to the weakest signals in the spectra of unmixed TTCF tryptic peptides shown in table 6.1. To illustrate the influence of ion suppression on peptide intensity, figure 6.5 shows the influence of ion suppression upon a single labelled peptide with an m/z value of 1135.5. The signal for this peptide following RP separation of the lysosomal autolysis products is compared to that obtained from the tryptic digest of TTCF and the mixture of labelled and unlabelled tryptic peptides before and after RP separation. From the results, it is clear that the peptide is weakly ionised in the tryptic digest of TTCF and is barely detectable in the unseparated mixture of labelled and unlabelled peptides. Following RP separation of this mixture, however, there is a 100-fold increase in its intensity as a result of the reduction in number of peptides co-eluting in the HPLC fraction. This leads to peptide 1135.5 being the dominant peptide present within the fraction. In contrast, following RP separation of the lysosomal autolysis products, over 100 peptides co-eluted with 1135.5. This led to a 10 fold reduction in its intensity compared to that observed

following RP separation of the less complex mixture of tryptic peptides. From these results, it is apparent that the influence of ion suppression upon the peptides is greatest in the more complex mixtures.

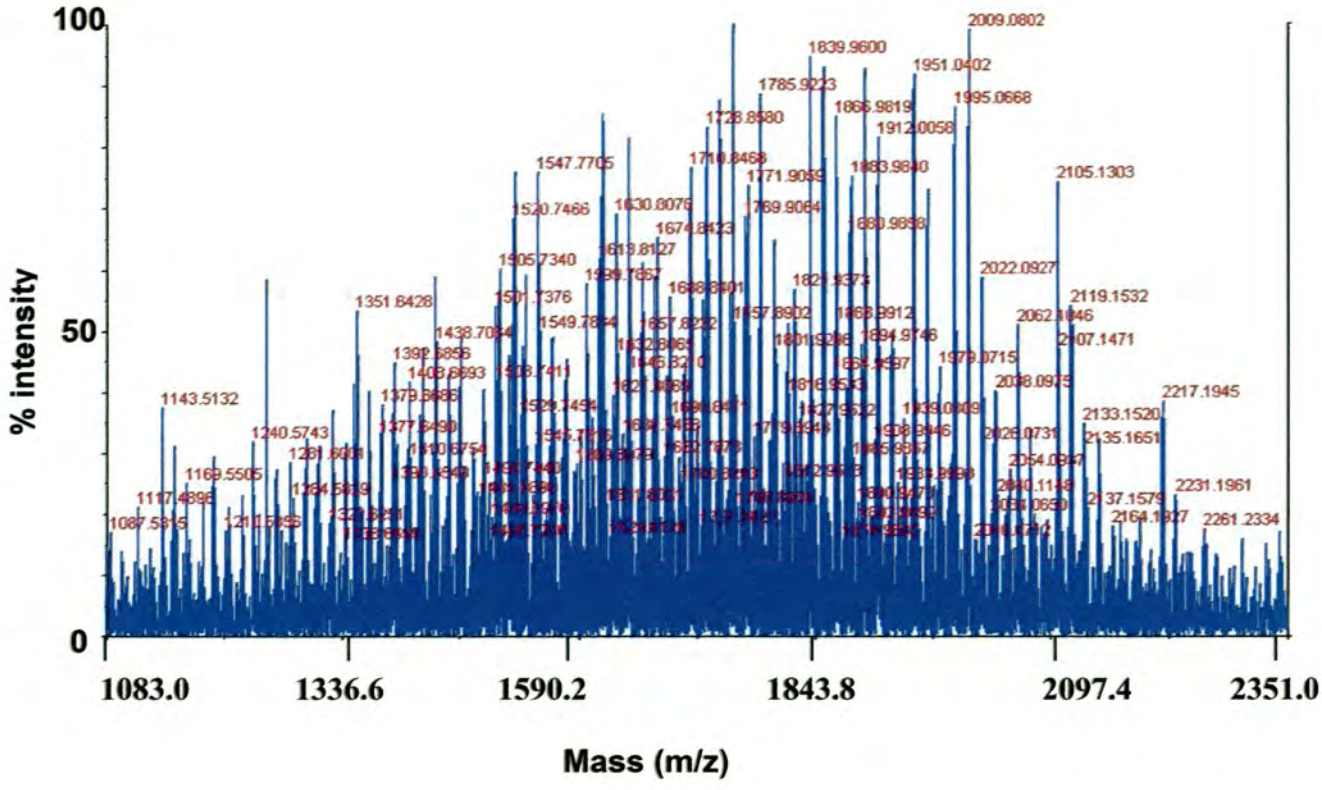


Figure 6.3 Identification of carbon-13 labelled characterised TTCF peptides in unseparated lysosomal autolysis products

Ten micrograms of lysosomes (+ 10mM DTT, 0.02% CHAPS) were digested at 37 ºC for 4 h and then the digestion products passed through a 3 KDa cut-off column. Ten pmoles of pre-digested carbon-13 labelled TTCF (same as that used in figure 6.1) was added to the lysosomal autolysis products and then a small proportion analysed by MALDI-TOF mass spectrometry.

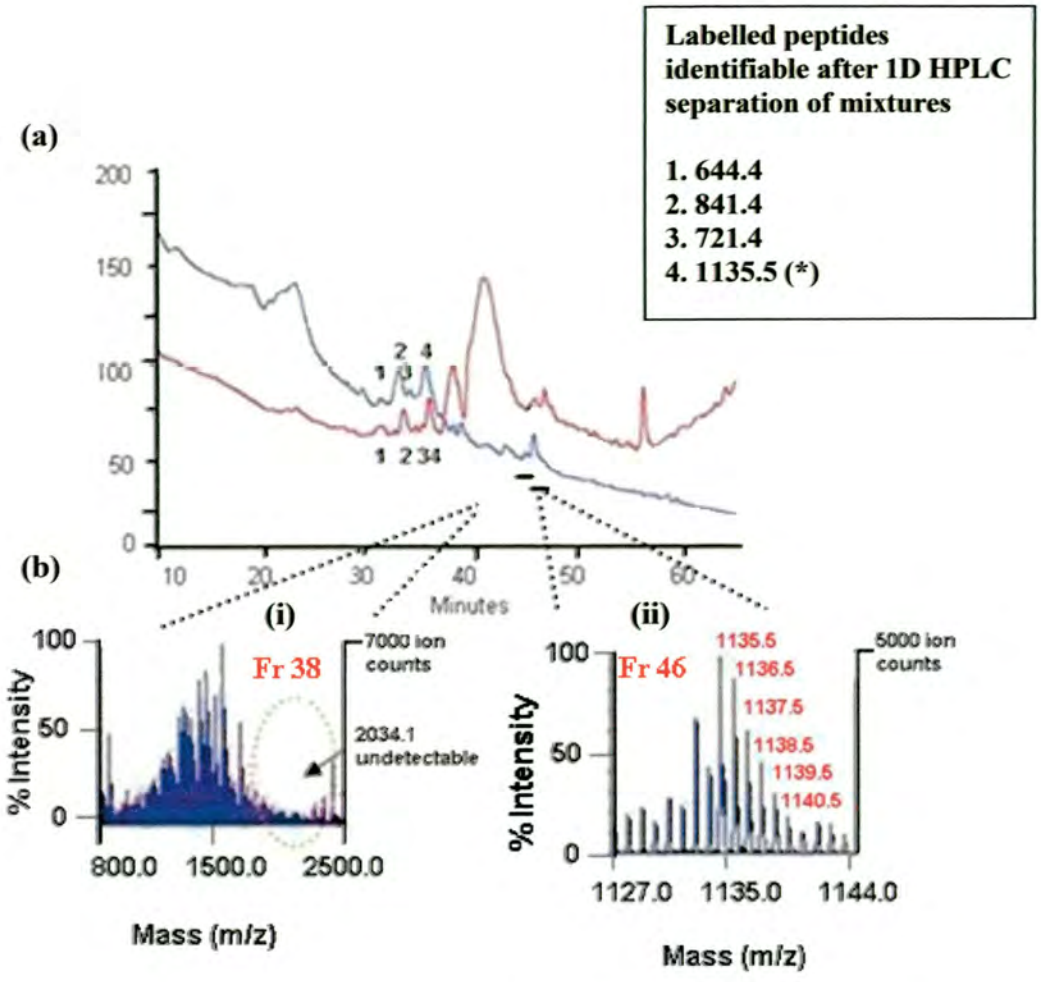


Figure 6.4 Identification of carbon-13 labelled TTCF peptides in lysosomal autolysis products following RP nano-flow LC separation

a) UV tracings from the RP separation of 10 pmoles of a carbon-13 labelled TTCF digest (same as one used in figure 6.1) either alone (red line) or mixed with lysosomal autolysis products (blue line) **b)** Mass spectrograms of 2 fractions (Fr 38 & Fr 46) from the RP separation of the mixture of TTCF peptides and lysosomal autolysis products. **(i)** The mass spectrogram of Fr 38 shows peptides of m/z value between 800-2500, with the labelled peptide of m/z valve 2034.1 not being detectable as a consequence of ion suppression. **(ii)** The mass spectrogram of Fr 46 is a blow up of the labelled TTCF peptide with an m/z value of 1135.5, which is shown to occur in this fraction as part of an overlapping peptide cluster.

Legend: (*) present but part of an overlap

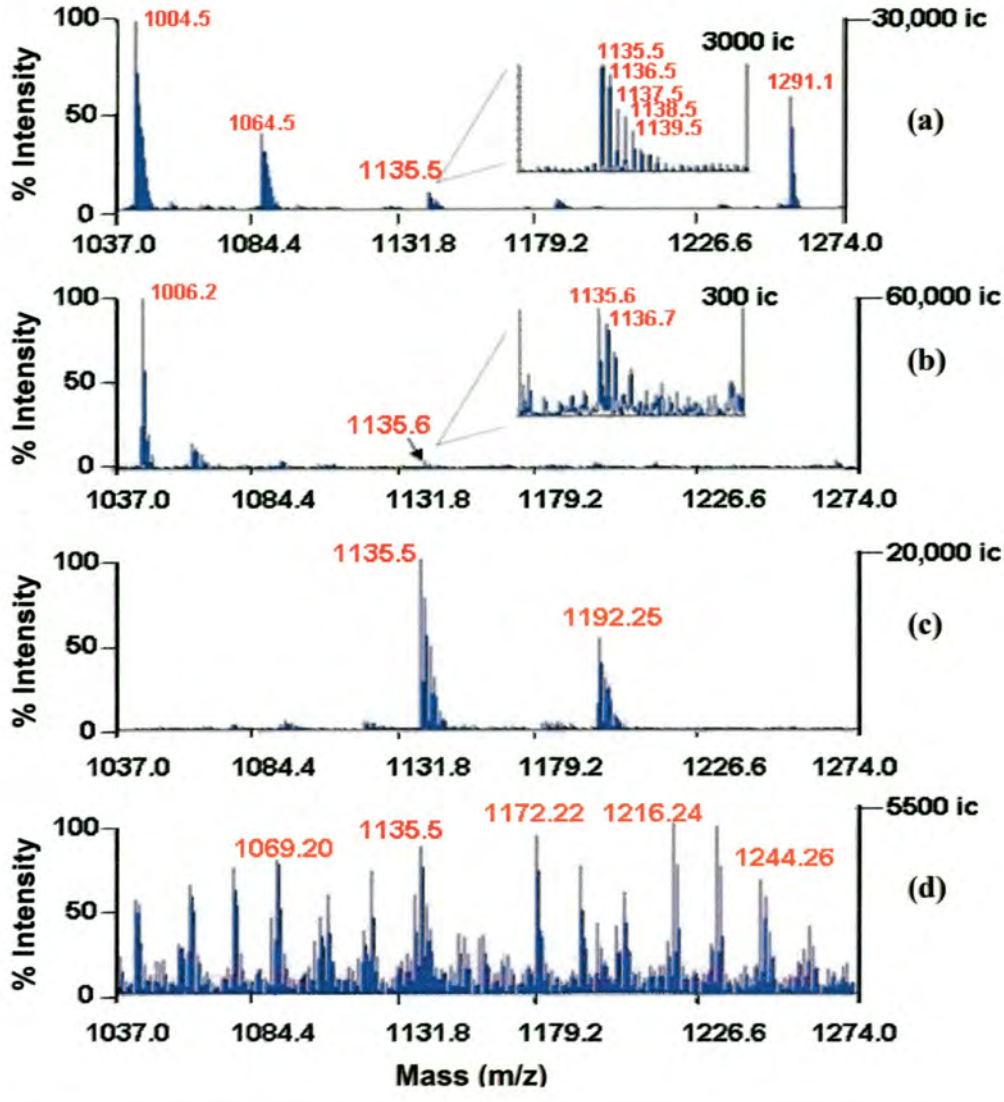


Figure 6.5 Influence of ion suppression upon the intensity of a labelled TTCF tryptic peptide

Signal intensity, as measured by the number of ion counts recorded for an ion during mass spectrometric data collection, was compared for the TTCF peptide of m/z value 1135.5 in different spectra. The spectra were collected from:

a) The unseparated tryptic digest of TTCF peptides alone
b) The unseparated mixture of labelled and unlabelled tryptic peptides (shown in figure 6.1b)
c) Following RP nano-flow LC separation of **(b)**
d) Following RP nano-flow LC separation of lysosomal autolysis products

Legend
ic: ion counts

6.4 Discussion

The experiments in this chapter were designed to examine the possibility that the lack of peptides in the previous set of experiments was simply a consequence of low sensitivity of the carbon-13 labelling technique. In these experiments, 10 pmoles of TTCF tryptic peptides were 'spiked' into mixtures of differing complexity at a relative low abundance and subsequently separated using RP chromatography. The results clearly show that detection in MALDI is initially dependant upon the complexity of samples. Thus, in the unseparated mixtures, the majority or all of the labelled TTCF peptides were undetectable as a result of ion suppression. The huge influence of ion suppression upon the intensity of peptide species in complex fractions was further demonstrated for the peptide with an m/z value of 1135. Importantly, it is also clear that the peptides that ionise less well and give rise to the lowest ion counts in uncomplex samples were most susceptible to becoming suppressed beyond the limits of detection in the complex HPLC fractions.

Following RP separation of the mixtures, all 8 of the TTCF peptides were identifiable following separation of the simpler peptide mixture generated from the digestion of 3 unlabelled proteins and 5 out of the 8 were detectable following RP separation of the more complex lysosomal autolysis mixture. To ensure that the identification of the labelled peptides was not biased due to the familiarity of the experiments with the TTCF tryptic fragments, an individual who was unfamiliar with the tryptic fragments was asked to identify the labelled peptides based solely upon their isotopic pattern. This person identified the labelled peptides in the fractions with ease.

The results considerably help the interpretation of the low abundance of TTCF fragments in lysosomal digests of intact TTCF that was described in the previous chapter. In those earlier experiments, much larger amounts of TTCF were used, with up to 500 μg of TTCF being digested. This is the equivalent of 10,000 pmoles of peptides following the complete digestion of TTCF. Since the limit of detection was shown to be below 10 pmoles in experiments in this study, other factors must play a role to explain why few labelled TTCF peptides were identified. Moreover, the experiments in the previous chapter employed the 2D separation

system which was earlier shown to improve separation and identification of labelled peptides compared to that of a 1D separation system.

Taken together, it was concluded from these results that the carbon-13 labelling technique had ample sensitivity to detect TTCF derived peptides among peptides recovered from lysosomal digests of TTCF, as described in the previous chapter. It is consequently concluded that small TTCF-derived peptides are extremely scarce in lysosomes. The possible reason for this observation will be discussed in the next chapter.

Chapter 7: Discussion

The experiments in this study were designed to test whether a novel ^{13}C labelling technique could be used to identify peptides of interest within complex mixtures. Such a labelling technique would have broad applicability in circumstances where it is desired to identify a protein and its breakdown products in complex samples recovered from biological systems. The spur to development has been the problem of distinguishing antigen-derived peptides in MHC class II or lysosome-extracts from antigen-pulsed antigen presenting cells. These peptides are of great interest in understanding how antigens are processed, and in rational vaccine design. With this in mind, the ultimate goal of this project was to use the isotope labelling technique to identify bacterial derived MHC class II associated peptides following pulsing experiments with antigen presenting cells and bacteria.

Biochemical analysis of complex peptide mixtures is both laborious and time consuming, often taking several months to complete. Isotopic labelling techniques have the potential to greatly simplify analysis by making peptides of interest distinguishable from other peptides within complex mixtures. Recently, using a somewhat different technique, the potential of stable isotope labelling was established by Meiring *et al.*, (2005). In that study, bacteria were cultured in minimal medium containing either $^{14}NH_4Cl$ or $^{15}NH_4Cl$, and then the outer membrane vesicles (OMV) containing Por A subsequently isolated and used to pulse dendritic cells in a 1:1 ratio. The MHC class II peptides were subsequently isolated, separated by HPLC and labelled peptides identified as 'spectral couplets' during mass spectrometric analysis. These spectral couplets arise because of a 1.2% average increase in mass in peptides labelled with ^{15}N. Using this technique, they identified 10 naturally processed and presented epitopes from the meningococcal outer membrane protein Por A. Two of these epitopes were found to be dominantly expressed at the cell surface and to be immunodominant $CD4^+T$ cell epitopes conserved between different strains. Such epitopes would therefore be suitable candidates in the design of epitope-based vaccines against infectious disease.

The technique examined in this study employed the stable isotope ^{13}C rather than ^{15}N. During development of the technique, ^{13}C was chosen in preference to ^{15}N

because the natural isotope pattern of peptides is dominated by that due to the content of ^{13}C, so a small change in ^{13}C content would discernibly alter the isotope overall pattern whereas much bigger changes in ^{15}N content would have been required. Indeed, Meiring *et al.*, (2005) had to make mixtures of 100% ^{14}N and ^{15}N to identify labelled peptides. Thus, carbon-13 labelling would be substantially cheaper when labelling large amounts of antigen. In addition, if the ratio of $^{14}NH_4Cl$ to $^{15}NH_4Cl$ labelled bacteria is not exactly 1:1 in pulsing experiments, this can lead to abundance differences between spectral couplets that can make analysis more difficult in very complex mixtures. With the ^{13}C technique, this is not an issue because the mass spectra are scanned for individual peptides that have a unique isotopic appearance to all other peptides within the spectra. However, other concerns do arise when using the ^{13}C labelling technique to identify labelled peptides in complex mixtures. Firstly, increased baseline spreading resulting from an increase in the number of labelled peaks for a peptide species leads to the occurrence of overlapping peptide species in more complex spectra. Secondly, the spreading of the overall ion count for a peptide over a greater number of isotopic peaks results in a decrease in mass spectrometric sensitivity. To minimise the influence of these two problems, the study opted to double the ^{13}C content of proteins.

Another important difference between the studies is the type of mass spectrometry used to characterise peptides. Meiring *et al.*, (2005) employed online nLC-ESI-MS in their study whereas this project employed MALDI-TOF spectrometric analysis. Online nLC-ESI-MS is becoming the more common method of characterisation of class II peptides. Tandem mass spectrometry provides both MS and MS/MS analysis of all peptides within the mixture and can detect sub-femtomole quantities of peptides. The high sensitivity of these mass spectrometers has led to the identification of class II peptides following nLC-ESI-MS of unfractionated class II samples (Nepom *et al.*, 2001; Meiring *et al.*, 2005). Carbon-13 labelling techniques can be successfully combined with instruments such as these. However, these instruments are more expensive to buy / maintain and are difficult to operate. Therefore, they are not widely accessible. In contrast, MALDI-TOF instruments are cheaper to maintain and far easier to operate, being more accessible to many research groups.

One of the aims stated at the beginning of the project was to feed whole bacteria to antigen presenting cells and subsequently identify T cell epitopes from immunogenic antigens. With this in mind, the study initially focussed upon work using *E. coli* outer membrane proteins. *E. coli* was chosen for use in this study for several reasons. Firstly, the genome of *E. coli* has recently been sequenced, providing invaluable information about the proteins encoded by the DNA sequence of the micro-organism (Blattner *et al.*, 1997). Secondly, *E. coli* is relatively safe to handle in the laboratory and is easy to grow on a large scale, producing satisfactory amounts of proteins for use in later experiments. Thirdly, *E. coli* is a pathogen, causing a variety of diseases ranging from gastroenteritis to urinary tract infections which are estimated to cause great medical and economic losses (Minino *et al.*, 2002). Therefore, the search for protective T cell epitopes conserved between different strains may prove to be beneficial in the design of vaccines. Indeed, the use of overlapping synthetic peptides has previously identified several murine T cell epitopes in OMP F, one of which was recognised in a genetically permissive manner by several MHC molecules (H-2 ^{d,k,b}). This epitope (OMP F $_{295\text{-}314}$) is in a transmembrane portion of OMP F and is conserved between different porins in *E. coli* (OMP C, Pho E) as well as between different species of Enterobacteriaceae, including *Salmonella enterica* serovar Typhimurium, *Salmonella enterica* serovar Typhi, and *Shigella flexneri*. It has also been shown to polarise the immune response in mice towards a T_H1 response that would be desirable against intracellular pathogens such as *Shigella* and *Salmonella*. Thus, the identification of genetically permissive T cell epitopes within *E. coli* outer membrane proteins may prove to be important in the design of peptide based vaccines.

For preliminary work, outer membrane extracts were chosen in preference to whole bacteria in order to simplify biochemical analysis during development and validation of the labelling technique. Initial experiments demonstrated that these outer membrane extracts contained a number of dominantly expressed proteins, many of which were bound by serum antibodies from a number of healthy individuals. These outer membrane proteins were therefore considered suitable candidates for ^{13}C labelling in order to investigate the presence of T cell epitopes. However, due to the complexity of these extracts, as demonstrated by enzyme

digestion followed by mass spectral analysis, the focus turned to validating the ^{13}C labelling technique in less complex mixtures. TTCF was selected because it was thought that the use of one antigen in experiments would make the identification of antigen-derived peptides an easier task and thus simplify biochemical analysis. Attention could then be refocused upon more complex groups of antigens such as the outer membrane proteins.

The study established that the ^{13}C technique was broadly applicable to labelling bacterial proteins. Labelling was demonstrated for native bacterial proteins, specifically numerous bacterial outer membrane proteins, and for foreign proteins expressed in *E. coli* in the form of TTCF. In each case, ^{13}C labelled peptides were readily distinguishable from their unlabelled counterparts on mass spectrometric analysis. Just a doubling of the ^{13}C content to give 2.3% labelling was shown to be sufficient to produce isotopically different peptides without leading to significant base line spreading and a decrease in sensitivity. Moreover the technique is very economical as only a small amount of ^{13}C enriched glucose is required to be added to minimal media in order to efficiently label a large amount of a desired protein. Together with previous data published from the laboratory, these results show that the ^{13}C labelling technique is applicable to a wide range of antigens that can be expressed either as recombinant proteins or extracted from bacteria grown in supplemented media (Zou *et al.*, 2004). Thus, many clinically important bacterial strains and bacterial toxins should be easily labelled using this technique. However, it may not be possible to label bacteria that require complex and undefined media to grow. For these pathogens, an alternative would be to choose bacterial antigens known to be important in virulence and produce labelled recombinant antigens rather than the whole organism. In addition, ^{13}C does not interfere with antigen processing and is very safe, making it a favourable alternative to radioactive labelling which is difficult to control on a large scale and to more discrete labels (e.g. biotin) which may become dissociated from the protein during processing or may interfere with processing of the molecule.

Initial experiments carried out using the outer membrane proteins demonstrated that identification of the ^{13}C-enriched peptides critically depended on the complexity of mass spectra during MALDI-TOF analysis because overlapping

peptides are abundant in complex spectra. In order to minimise / prevent the occurrence of overlapping peptides, a 2D IEX-RP nano-flow LC system was set up in the laboratory and used to separate an enzymatic digest of *E. coli* outer membrane proteins. Separation of the digest by the 2D IEX-RP nano-flow LC system led to a reduction in the occurrence of overlapping peptides and an increase in the number of detectable peptide species compared to 1D RP nano-flow separation. This was thought to be due to a decrease in the number of peptides co-eluting at the same retention time, and a consequent reduction in ion suppressive effects within these fractions. It is clear that multidimensional separation techniques will be critical in applying ^{13}C-enrichment to identifying antigen derived peptides in class II peptide pools if MALDI-TOF MS is used. Other ionisation techniques are less susceptible to ion suppression effects, and may be better partners to ^{13}C labelling.

It was decided to apply the techniques first to an analysis of the peptide intermediates generated within lysosomes. These peptides are interesting because they would identify areas of the antigen that are released during antigenic processing for binding to MHC class II molecules and presentation to T cells. Furthermore, these experiments would also validate the use of the ^{13}C labelling technique in a complex biological system before going onto to tackle more complicated studies analysing class II eluted peptides. Moreover, it was thought that this would be a less challenging experiment than analysis of HLA class II-bound peptides, because it should be possible to ensure that labelled antigen is one of the most prevalent proteins in the samples. Surprisingly, few ^{13}C labelled peptides were identified following 2D IEX-RP nano-flow HPLC separation of lysosomal processing products (<3KDa) following the in-vitro digestion of TTCF. This was the case even when large amounts of antigen were digested by much smaller amounts of lysosomes (5:1 ratio), and in spite of laborious and time consuming analysis. Only two peptides were found that had isotope patterns indicative of their being derived from ^{13}C-labelled TTCF.

It was a concern that the scarcity of labelled processing products may be a consequence of deficiencies in the labelling and detection systems, so the next set of experiments were set up to define the weaknesses and sensitivity of the technique. The principal weaknesses were that overlapping peptides occurred frequently in

complex spectra and that many peptides with labelled isotopic appearances gave rise to very low ion counts. Although many of these weak species had isotopic patterns resembling labelled peptides, many unlabelled peptides can have a labelled appearance at lower ion counts. As these fractions are quickly exhausted, it was impossible to determine whether these species are in fact labelled. Further experiments demonstrated that the techniques as they were applied had adequate sensitivity to detect TTCF peptides in lysosomal digests, had they been present, as much smaller molar quantities of pre-digested TTCF peptides (as small as ~10 pmoles) added to lysosomes were detected with efficiencies >50% employing a 1D separation strategy. These results demonstrated that the ^{13}C labelling technique combined with 1D or 2D depending on complexity is suitable for detecting small levels of peptides in complex mixtures.

Therefore, the results demonstrate that TTCF processing products are extremely scarce in lysosomes. It has been previously demonstrated that following digestion of radio-labelled TTCF in B cells, radiolabelled fragments are only recovered from areas protected by bound immunoglobulin, with non-protected parts being rapidly destroyed by lysosomal enzymes. In this study, however, in-vitro digestion generated five larger digestion fragments detectable by SDS-PAGE that persisted till the end of the digestion period despite a change in digestion parameters (pH) or processing compartments used in the experiments. As a consequence, it was thought that one reason for the lack of digestion species may have been due to the stability of larger digestion products during in-vitro digestion and the lack of production of smaller digestion products from the large digestion species. This may be due to factors which are present in cells but absent from the lysosomal compartments used during in-vitro experiments. These may include chaperones that aid in protein unfolding. Alternatively, a proportion of these larger fragments may have been digested into smaller peptides that were subsequently destroyed by lysosomal enzymes.

Previous to this work, there had been publications that had indicated that TTCF derived peptides should be detectable following in-vitro lysosomal digestion. Firstly, Hewitt, *et al.*, (1997) demonstrated that following the unfolding of TTCF, there are many cut sites for the enzymes cathepsin D and E within the antigen. In-

vitro digestion with both of these enzymes generated many species which were detected using mass spectrometric analysis (Hewitt *et al.*, 1997). Secondly, in-vitro lysosomal processing has been used successfully in the identification of digestion products (<3KD) of myelin basic protein (MBP) (Beck *et al.*, 2001; Burster *et al.*, 2004). It was shown that large fragments of MBP are initially detectable by reverse phase HPLC. Over a 6hr digestion period, these larger fragments are subsequently degraded into a number of smaller digestion products between 20-40 amino acids in length detectable by mass spectrometry, before being finally degraded into peptides of only a few amino acids in length and single amino acids. An important difference may be that both Hewitt et al (1997) and Beck *et al.*, (2001) used denatured TTCF and MBP respectively in digestion experiments. In contrast, TTCF was not denatured prior to digestion in this study. It is known that denaturation in the acidic endocytic pathway is required for efficient processing of antigens and occurs in the presence of reducing enzymes such as GILT (Jensen, 1993; Watts, 2001). It may be the case that if TTCF had been heat denatured prior to digestion, then the large fragments may have been digested into smaller fragments detectable by mass spectrometric analysis. This would indicate that another factor missing from our digestion experiments might have been a strong denaturing environment.

Based upon the results gained from this study, there are modifications that could be made to improve the experiment. As already mentioned, TTCF could be heat denatured prior to digestion to determine whether antigen unfolding would lead to the generation of smaller digestion species detectable by mass spectrometric analysis. In addition, the results in this study indicate that isotopic analysis alone can not be used to identify peptides unambiguously. Therefore, in future experiments, the sample should be split and MS/MS analysis carried out on peptides identified with a 'labelled' appearance to confirm their identity. To determine whether the ^{13}C technique is suitable for identifying naturally processed and presented peptides, the ^{13}C labelling technique could be used to identify naturally processed and presented peptides from an antigen that has already been studied in the way. This is currently underway in the laboratory, with experiments being carried out to identify naturally processed and presented peptides from α3 (IV) which has been previously identified by biochemical techniques (Phelps *et al.*, 1996).

Analysis of lysosomal peptides has illuminated TTCF processing but cast no light on the TTCF peptides that are presented to T cells. Unfortunately, there was insufficient time to examine ^{13}C-labelled TTCF-derived peptides eluted from HLA class II molecules. However, the preliminary results from this project indicate that it is very likely that application of the technique will allow identification of TTCF peptides among peptides eluted from class II molecules of TTCF-pulsed APC. Secondly, this approach would take into consideration the effects of MHC guided processing upon antigen processing and T cell epitope production which the in-vitro digestion system used in these experiments did not take into consideration (Sercarz *et al.*, 1993; Sercarz and Maverakis, 2003). In addition, in-vitro processing does not take into consideration the influence of the B cell receptor which has been shown to sculpt antigen processing of TTCF following internalisation of the antigen-receptor complex (Watts *et al.*, 1998). Therefore, a better system would be to use TTCF specific B cell lines and isolate class II associated peptides following internalisation, processing and presentation of TTCF.

Despite the discussed limitations, the ^{13}C labelling technique shows great potential in identifying naturally processed and presented T cell epitopes and could be used to label potential vaccine candidates for the identification of T cell epitopes. This may be desirable for antigens, such as the Por A protein from meningococcal species, which have different immunogenicities between different strains. In addition, it may be desirable to identify T cell epitopes that are conserved between different species of bacteria, which is the case for the OMP F T cell epitope. For these antigens, identifying immunogenic peptide sequences that are conserved between different strains and that lead to the generation of a protective immune response may prove beneficial in vaccine design. There are currently a number of methods employed to identify potential vaccine candidates. The most recent advents in this area are reverse vaccinology and 2D gel electrophoresis (Nilsson, 2002; Adu-Bobie *et al.*, 2003; Capecchi *et al.*, 2004). If desired, potential vaccine candidates identified by these methods could subsequently be ^{13}C labelled and potential T cell epitopes identified by examining peptides bound to class II molecules. Alternatively, the ^{13}C labelling technique could be used to label whole bacteria and then naturally processed and presented T cell epitopes characterised in order to identify

immunogenic bacterial proteins. The identification of naturally processed and presented peptides from whole bacteria has only been carried to date for $CD8^+$ T cells. However, advances in mass spectrometric sensitivity may allow these techniques to successfully be used to identify $CD4^+$ T cell epitopes. This technique would be favourable because it would identify proteins that are presented to T cells. This would select out those proteins that are of immunological relevance rather than identifying many proteins from bacteria that are not presented to T cells.

Appendix 1

Sequences and molecular weights of the peptides used in HPLC development

Peptides purchased from PerSeptive Biosystems as part of a Peptide Standards Kit (calibration mixture 2).

Peptide	Sequence	Molecular weight with +1 charge ($M+H^+$)
Angiotensin I	DRVYIHPFHL	1296.68
ACTH		
•1-17 clip	SYSMEHFRWGKPVGRKR	2093.08
•18-39 clip	RPVKVYTNGVEEESSEAFPSEM	2465.19
•7-38 clip	FRWGKPVGRKRRPVKVYTNGVEEESSEAFPSE	3657.92
Insulin, Bovine	MALWTRLRPLLALLALWPPPPARAFVNQHLCG SHLVEALYLVCGERGFFYTPKARREVEGPQVG ALELAGGPGAGGLEGPPQKRGIVEQCCASVCS LYQLENYCN	5730.60

Amino acid code

G	Glycine	**P**	Proline
A	Alanine	**V**	Valine
L	Leucine	**I**	Isoleucine
M	Methionine	**C**	Cysteine
F	Phenylalanine	**Y**	Cyrosine
W	Tryptophan	**H**	Histidine
K	Lysine	**R**	Arginine
Q	Glutamine	**N**	Asparagine
E	Glutamic acid	**D**	Aspartic acid
S	Serine	**T**	Threonine

Appendix 2

Sequences of the outer membrane proteins matched by experimental peptide masses generated by trypsin digestion of protein bands 1, 2 and 3

Omp A (band 3)

- MOWSE score: **107**
- Number of mass matches searched: **39**
- Number of mass values matched: **11**
- Sequence coverage: **55%**

Matched peptides shown in BOLD RED

```
  1 MKKTAIAIAV ALAGFATVAQ AAPKDNTWYT GAKLGWSQYH DTGFINNNGP
 51 THENQLGAGA FGGYQVNPYV GFEMGYDWLG RMPYKGSVEN GAYKAQGVQL
101 TAKLGYPITD DLDVYTRLGG MVWRADTKSN VYGKNHDTGV SPVFAGGVEY
151 AITPEIATRL EYQWTNNIGD AHTIGTRPDN GMLSLGVSYR FGQGEAAPVV
201 APAPAPAPEV QTKHFTLKSD VLFTFNKATL KPEGQAALDQ LYSQLSNLDP
251 KDGSVVVLGY TDRIGSDAYN QALSERRAQS VVDYLISKGI PADKISARGM
301 GESNPVTGNT CDNVKQRAAL IDCLAPDRRV EIEVKGIKDV VTQPQA
```

Start - end	Mr (experimental)	Mr (calculated)	Sequence
4-33	2990.44	2990.59	K.TAIAIAVALAGFATVAQAAPKAAP KDNTWYT GAK.L
25-33	1054.47	1054.47	K.DNTWYT GAK.L
104-117	1639.81	1639.81	K.LGYPITD DLDVYTR .L
118-124	817.45	817.43	R.LGG MVWR.A
135-159	2600.27	2600.29	K.NHDTGV SPVFAGGVEY AITPEIATR.L
160-190	3477.61	3477.67	R.LEYQWTNNIGDAHTIGTRPDNGML SLGVSYR.F
191-213	2231.15	2231.16	R.FGQGEAAPVVAPAPAPAPEVQTK. H
228-251	2599.31	2599.35	K.ATL KPEGQAALDQ LYSQLSNLDPK.D
252-263	1279.63	1279.64	K.DGSVVVLGY TDR.I
264-276	1422.67	1422.67	R.IGSDAYN QALSER.R
277-288	1377.77	1377.76	R.RAQS VVDYLISK.G

Omp C (band 2)

- MOWSE score: **181**
- Number of mass matches searched: **53**
- Number of mass values matched: **18**
- Sequence coverage: **66%**

Matched peptides shown in BOLD RED

```
  1 MKVKVLSLLV PALLVAGAAN AAEVYNKDGN KLDLYGKVDG LHYFSDDKSV
 51 DGDQTYMRLG FKGETQVTDQ LTGYGQWEYQ IQGNSAENEN NSWTRVAFAG
101 LKFQDVGSFD YGRNYGVVYD VTSWTDVLPE FGGDTYGSDN FMQQRGNGFA
151 TYRNTDFFGL VDGLNFAVQY QGKNGSVSEG MTNNGREALR QNGDGVGGSI
201 TYDYEGFGIG AAVSSSKRTD DQNSPLYIGN GDRAETYTGG LKYDANNIYL
251 AAQYTQTYNA TRVGSLGWAN KAQNFEAVAQ YQFDFGLRPS LAYLQSKGKN
301 LGVINGRNYD DEDILKYVDV GATYYFNKNM STYVDYKINL LDDNQFTRDA
351 GINTDNIVAL GLFYQF<
```

Start - end	Mr (experimental)	Mr (calculated)	Sequence
28-37	1121.56	1121.57	K.DGNKLDLYGK.V
38-48	1294.59	1294.58	K.VDGLHYFSDDK.S
38-58	2447.08	2447.07	K.VDGLHYFSDDKSVDGDQTYMR.L
49-58	1170.51	1170.51	K.SVDGDQTYMR.L
63-95	3802.67	3802.68	K.GETQVTDQLTGYGQWEYQIQGNSAENENNSWTR.V
96-102	704.41	704.42	R.VAFAGLK.F
103-113	1289.55	1289.57	K.FQDVGSFDYGR.N
114-145	3659.65	3659.61	R.NYGVVYDVTSWTDVLPEFGGDTYGSDNFMQQR.G
146-153	884.41	884.41	R.GNGFATYR.N
154-173	2232.09	2232.09	R.NTDFFGLVDGLNFAVQYQGK.N
218-233	1819.81	1819.84	K.RTD DQNSPLYIGN GDR.A
219-233	1663.78	1663.74	R.TD DQNSPLYIGN GDR.A
243-262	2353.10	2353.10	K.YDANNIYLAAQYTQTYNATR.V
263-271	930.49	930.49	R.VGSLGWANK.A
272-297	2990.50	2990.49	K.AQNFEAVAQYQFDFGLRPSLAYLQSK.G
300-307	841.50	841.48	K.NLGVINGR.N
317-328	1438.67	1438.68	K.YVDVGATYYFNK.N
338-348	1347.68	1347.68	K.INLLDDNQFTR.D

Omp F (band 1)

- MOWSE score: **113**
- Number of mass values searched: **50**
- Number of mass values matched: **14**
- Sequence coverage: **50%**

Matched peptides shown in BOLD RED:

1 MMKRNILAVI VPALLVAGTA NAAEIYNK**DG NKVDLYGKAV GLHYFSKGNG**
51 **ENSYGGNGDM TYAR**LGFKGE TQINSDLTGY GQWEYNFQGNNSEGADAQTG
101NKTR**LAFAGL KYADVGSFDY GR**NYGVVYDA LGYTDMLPEFGGDTAYSDDF
151FVGR**VGGVAT YRNSNFFGLV DGLNFAVQYL GK**NERDTARR**SNGDGVGGSI**
201**SYEYEGFGIV GAYGAADRTN LOEESSLGKG** KK**AEOWATGLKYDANNIYLA**
251**ANYGETR**NAT PITNK**FTNTS GFANKTODVL LVAOYOFDFGLRPSIAYTK**S
301KAKDVEGIGD VDLVNYFEVG ATYYFNKNMS TYVDYIINQIDSDNKLGVGS
351DDTVAVGIVY QF

Start - end	Mr (experimental)	Mr (calculated)	Sequence
29-38	1107.53	1107.56	K.DGNK VDLYGKA
39-47	1020.54	1020.54	K.AVGLHYFSK.G
48-64	1761.76	1761.70	K.GNGENSYGGNGDMTYAR.L
105-111	718.44	718.44	R.LAFAGLK.Y
112-122	1248.53	1248.54	K.YADVGSFDYGR.N
155-162	821.46	821.44	R.VGGVATYR.N
163-182	2202.13	2202.11	R.NSNFFGLV DGLNFAVQYLGK.N
191-218	2767.23	2767.24	R.SNGDGVGGSISYEYEGFGIV GAYGAAR.T
219-229	1204.55	1204.59	R.TNLOEESSLGK.G
232-241	1130.60	1130.61	K.KAEQWATGL K.Y
233-241	1002.51	1002.51	K.AEOWATGL K.Y
242-257	1846.86	1846.85	K.YDANNIYLAANYGETR.N
266-275	1085.51	1085.51	K.FTNTSGFANK.T
276-299	2772.42	2772.45	K.TODVLLVAOYQYQFDFGLRPSI AYTK.S

Bibliography

Ackerman, A. L. and P. Cresswell (2004). Cellular mechanisms governing cross-presentation of exogenous antigens. *Nat Immunol* **5** (7): 678-84.

Aderem, A. and D. M. Underhill (1999). Mechanisms of phagocytosis in macrophages. *Annu Rev Immunol* **17**: 593-623.

Aebersold, R. and M. Mann (2003). Mass spectrometry-based proteomics. *Nature* **422** (6928): 198-207.

Al-Attiyah, R., F. A. Shaban, H. G. Wiker, F. Oftung and A. S. Mustafa (2003). Synthetic peptides identify promiscuous human Th1 cell epitopes of the secreted mycobacterial antigen MPB70. *Infect Immun* **71** (4): 1953-60.

Alfonso, C., G. S. Williams, J. O. Han, J. A. Westberg, O. Winqvist and L. Karlsson (2003). Analysis of H2-O influence on antigen presentation by B cells. *J Immunol* **171** (5): 2331-7.

Amigorena, S. and C. Bonnerot (1998). Role of B-cell and Fc receptors in the selection of T-cell epitopes. *Curr Opin Immunol* **10** (1): 88-92.

Amigorena, S. and C. Bonnerot (1999). Fc receptors for IgG and antigen presentation on MHC class I and class II molecules. *Semin Immunol* **11** (6): 385-90.

Amigorena, S., J. R. Drake, P. Webster and I. Mellman (1994). Transient accumulation of new class II MHC molecules in a novel endocytic compartment in B lymphocytes. *Nature* **369** (6476): 113-20.

Amigorena, S., D. Lankar, V. Briken, L. Gapin, M. Viguier and C. Bonnerot (1998). Type II and III receptors for immunoglobulin G (IgG) control the presentation of different T cell epitopes from single IgG-complexed antigens. *J Exp Med* **187** (4): 505-15.

Antoniou, A. N., S. L. Blackwood, D. Mazzeo and C. Watts (2000). Control of antigen presentation by a single protease cleavage site. *Immunity* **12** (4): 391-8.

Antoniou, A. N. and C. Watts (2002). Antibody modulation of antigen presentation: positive and negative effects on presentation of the tetanus toxin antigen via the murine B cell isoform of FcgammaRII. *Eur J Immunol* **32** (2): 530-40.

Arunachalam, B., U. T. Phan, H. J. Geuze and P. Cresswell (2000). Enzymatic reduction of disulfide bonds in lysosomes: characterization of a gamma-interferon-inducible lysosomal thiol reductase (GILT). *Proc Natl Acad Sci U S A* **97** (2): 745-50.

Babbitt, B. P., P. M. Allen, G. Matsueda, E. Haber and E. R. Unanue (1985). Binding of immunogenic peptides to Ia histocompatibility molecules. *Nature* **317** (6035): 359-61.

Babbitt, B. P., P. M. Allen, G. Matsueda, E. Haber and E. R. Unanue (2005). Binding of immunogenic peptides to Ia histocompatibility molecules. 1985. *J Immunol* **175** (7): 4163-5.

Bakke, O. and B. Dobberstein (1990). MHC class II-associated invariant chain contains a sorting signal for endosomal compartments. *Cell* **63** (4): 707-16.

Balaji, K. N. and W. H. Boom (1998). Processing of Mycobacterium tuberculosis bacilli by human monocytes for CD4+ alphabeta and gammadelta T cells: role of particulate antigen. *Infect Immun* **66** (1): 98-106.

Barnes, K. A. and R. N. Mitchell (1995). Detection of functional class II-associated antigen: role of a low density endosomal compartment in antigen processing. *J Exp Med* **181** (5): 1715-27.

Beck, H., G. Schwarz, C. J. Schroter, M. Deeg, D. Baier, S. Stevanovic, E. Weber, C. Driessen and H. Kalbacher (2001). Cathepsin S and an asparagine-specific endoprotease dominate the proteolytic processing of human myelin basic protein in vitro. *Eur J Immunol* **31** (12): 3726-36.

Beck, S. and J. Trowsdale (2000). The human major histocompatability complex: lessons from the DNA sequence. *Annu Rev Genomics Hum Genet* **1**: 117-37.

Beers, C., A. Burich, M. J. Kleijmeer, J. M. Griffith, P. Wong and A. Y. Rudensky (2005). Cathepsin S controls MHC class II-mediated antigen presentation by epithelial cells in vivo. *J Immunol* **174** (3): 1205-12.

Bjorkman, P. J., M. A. Saper, B. Samraoui, W. S. Bennett, J. L. Strominger and D. C. Wiley (1987). Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* **329** (6139): 506-12.

Blattner, F.R., G. Plunkett, C.A. Bloch, N.T. Perna, V. Burland, M. Riley, J. Collado-Vides, J.D. Glasner, C.K. Rode, G.F. Mayhew, J. Gregor, N.W. Davis, H.A. Kirkpatrick, M.A. Goeden, D.J. Rose, B. Mau, and Shao Y (1997). The complete genome sequence of *Escherichia coli K-12*. *Science* **277** (5331): 1453-74.

Blum, J. S. and P. Cresswell (1988). Role for intracellular proteases in the processing and transport of class II HLA antigens. *Proc Natl Acad Sci U S A* **85** (11): 3975-9.

Bolin, C. A. and A. E. Jensen (1987). Passive immunization with antibodies against iron-regulated outer membrane proteins protects turkeys from *Escherichia coli* septicemia. *Infect Immun* **55** (5): 1239-42.

Bonnerot, C., D. Lankar, D. Hanau, D. Spehner, J. Davoust, J. Salamero and W. H. Fridman (1995). Role of B cell receptor Ig alpha and Ig beta subunits in MHC class II-restricted antigen presentation. *Immunity* **3** (3): 335-47.

Braud, V. M., D. S. Allan and A. J. McMichael (1999). Functions of nonclassical MHC and non-MHC-encoded class I molecules. *Curr Opin Immunol* **11** (1): 100-8.

Brocke, P., E. Armandola, N. Garbi and G. J. Hammerling (2003). Downmodulation of antigen presentation by H2-O in B cell lines and primary B lymphocytes. *Eur J Immunol* **33** (2): 411-21.

Brown, E. J. (1995). Phagocytosis. *Bioessays* **17** (2): 109-17.

Brown, J. H., T. S. Jardetzky, J. C. Gorga, L. J. Stern, R. G. Urban, J. L. Strominger and D. C. Wiley (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364** (6432): 33-9.

Bryant, P. and H. Ploegh (2004). Class II MHC peptide loading by the professionals. *Curr Opin Immunol* **16** (1): 96-102.

Burster, T., A. Beck, E. Tolosa, V. Marin-Esteban, O. Rotzschke, K. Falk, A. Lautwein, M. Reich, J. Brandenburg, G. Schwarz, H. Wiendl, A. Melms, R. Lehmann, S. Stevanovic, H. Kalbacher and C. Driessen (2004). Cathepsin G, and not the asparagine-specific endoprotease, controls the processing of myelin basic protein in lysosomes from human B lymphocytes. *J Immunol* **172** (9): 5495-503.

Busch, D. H., H. G. Bouwer, D. Hinrichs and E. G. Pamer (1997). A nonamer peptide derived from Listeria monocytogenes metalloprotease is presented to cytolytic T lymphocytes. *Infect Immun* **65** (12): 5326-9.

Busch, R., I. Cloutier, R. P. Sekaly and G. J. Hammerling (1996). Invariant chain protects class II histocompatibility antigens from binding intact polypeptides in the endoplasmic reticulum. *Embo J* **15** (2): 418-28.

Busch, R., G. Strang, K. Howland and J. B. Rothbard (1990). Degenerate binding of immunogenic peptides to HLA-DR proteins on B cell surfaces. *Int Immunol* **2** (5): 443-51.

Caccamo, N., S. Meraviglia, C. La Mendola, S. Bosze, F. Hudecz, J. Ivanyi, F. Dieli and A. Salerno (2004). Characterization of HLA-DR- and TCR-binding residues of an immunodominant and genetically permissive peptide of the 16-kDa protein of Mycobacterium tuberculosis. *Eur J Immunol* **34** (8): 2220-9.

Carrasco-Marin, E., J. E. Paz-Miguel, P. Lopez-Mato, C. Alvarez-Dominguez and F. Leyva-Cobian (1998). Oxidation of defined antigens allows protein unfolding and increases both proteolytic processing and exposes peptide epitopes which are recognized by specific T cells. *Immunology* **95** (3): 314-21.

Casciola-Rosen, L. A. and A. L. Hubbard (1991). "Hydrolases in intracellular compartments of rat liver cells. Evidence for selective activation and/or delivery." J *Biol Chem* **266** (7): 4341-7.

Castellino, F. and R. N. Germain (1995). Extensive trafficking of MHC class II-invariant chain complexes in the endocytic pathway and appearance of peptide-loaded class II in multiple compartments. *Immunity* **2** (1): 73-88.

Castellino, F., F. Zappacosta, J. E. Coligan and R. N. Germain (1998). Large protein fragments as substrates for endocytic antigen capture by MHC class II molecules. *J Immunol* **161** (8): 4048-57.

Cella, M., A. Engering, V. Pinet, J. Pieters and A. Lanzavecchia (1997). Inflammatory stimuli induce accumulation of MHC class II complexes on dendritic cells. *Nature* **388** (6644): 782-7.

Chapman, H. A. (1998). Endosomal proteolysis and MHC class II function. *Curr Opin Immunol* **10** (1): 93-102.

Chart, H. and E. Griffiths (1985). Antigenic alteration of contaminating lipopolysaccharide during extraction of *Escherichia coli* outer-membrane proteins from polyacrylamide gels. *J Gen Microbiol* **131** (10): 2847-50.

Chen, X., L. M. Smith and E. M. Bradbury (2000). Site-specific mass tagging with stable isotopes in proteins for accurate and efficient protein identification. *Anal Chem* **72** (6): 1134-43.

Chesnut, R. W., S. M. Colon and H. M. Grey (1982). Antigen presentation by normal B cells, B cell tumors, and macrophages: functional and biochemical comparison. *J Immunol* **128** (4): 1764-8.

Chicz, R. M., R. G. Urban, W. S. Lane, J. C. Gorga, L. J. Stern, D. A. Vignali and J. L. Strominger (1992). Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature* **358** (6389): 764-8.

Chicz, R. M., R. G. Urban, J. C. Gorga, D. A. Vignali, W. S. Lane and J. L. Strominger (1993). Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. *J Exp Med* **178** (1): 27-47.

Chu, C. T. and S. V. Pizzo (1993). Receptor-mediated antigen delivery into macrophages. Complexing antigen to alpha 2-macroglobulin enhances presentation to T cells. *J Immunol* **150** (1): 48-58.

Collins, D. S., E. R. Unanue and C. V. Harding (1991). Reduction of disulfide bonds within lysosomes is a key step in antigen processing. *J Immunol* **147** (12): 4054-9.

Corr, M., L. F. Boyd, E. A. Padlan and D. H. Margulies (1993). H-2Dd exploits a four residue peptide binding motif. *J Exp Med* **178** (6): 1877-92.

Cresswell, P. (1994). Assembly, transport, and function of MHC class II molecules. *Annu Rev Immunol* **12**: 259-93.

Davidson, H. W., P.A. Reid, A. Lanzavecchia, and C. Watts (1991). "Processed antigen binds to newly synthesized MHC class II molecules in antigen-specific B lymphocytes." Cell **67** (1): 105-16.

Davidson, H. W. and C. Watts (1989). Epitope-directed processing of specific antigen by B lymphocytes. *J Cell Biol* **109** (1): 85-92.

Davidson, H. W., M. A. West, and C. Watts (1990). Endocytosis, intracellular trafficking, and processing of membrane IgG and monovalent antigen/membrane IgG complexes in B lymphocytes. *J Immunol* **144** (11): 4101-9.

de Boer, E. C., R. F. Bevers, K. H. Kurth and D. H. Schamhart (1996). Double fluorescent flow cytometric assessment of bacterial internalization and binding by epithelial cells. *Cytometry* **25** (4): 381-7.

Delvig, A. A. and J. H. Robinson (1998). Different endosomal proteolysis requirements for antigen processing of two T-cell epitopes of the M5 protein from viable Streptococcus pyogenes. *J Biol Chem* **273** (6): 3291-5.

Delvig, A. A. and J. H. Robinson (1998). Two T cell epitopes from the M5 protein of viable Streptococcus pyogenes engage different pathways of bacterial antigen processing in mouse macrophages. *J Immunol* **160** (11): 5267-72.

Demotz, S., H. M. Grey, E. Appella and A. Sette (1989). Characterization of a naturally processed MHC class II-restricted T-cell determinant of hen egg lysozyme. *Nature* **342** (6250): 682-4.

Demotz, S., A. Lanzavccchia, U. Eisel, H. Nielmann, C. Widmann, and G Corradin (1989). Delineation of several DR-restricted tetanus toxin T cell epitopes. *J Immunol* **142** (2): 394-402.

Dengjel, J., P. Decker, O. Schoor, F. Altenberend, T. Weinschenk, H. G. Rammensee and S. Stevanovic (2004). Identification of a naturally processed cyclin D1 T-helper epitope by a novel combination of HLA class II targeting and differential mass spectrometry. *Eur J Immunol* **34** (12): 3644-51.

Denzin, L. K. and P. Cresswell (1995). HLA-DM induces CLIP dissociation from MHC class II alpha beta dimers and facilitates peptide loading. *Cell* **82**(1): 155-65.

Denzin, L. K., D. B. Sant'Angelo, C. Hammond, M. J. Surman and P. Cresswell (1997). Negative regulation by HLA-DO of MHC class II-restricted antigen processing. *Science* **278** (5335): 106-9.

Deussing, J., W. Roth, P. Saftig, C. Peters, H. L. Ploegh and J. A. Villadangos (1998). Cathepsins B and D are dispensable for major histocompatibility complex class II-mediated antigen presentation. *Proc Natl Acad Sci U S A* **95** (8): 4516-21.

DiBrino, M., K. C. Parker, J. Shiloach, M. Knierman, J. Lukszo, R. V. Turner, W. E. Biddison and J. E. Coligan (1993). Endogenous peptides bound to HLA-A3 possess a specific combination of anchor residues that permit identification of potential antigenic peptides. *Proc Natl Acad Sci U S A* **90** (4): 1508-12.

Diethelm-Okita, B. M., D. K. Okita, L. Banaszak and B. M. Conti-Fine (2000). Universal epitopes for human CD4+ cells on tetanus and diphtheria toxins. *J Infect Dis* **181** (3): 1001-9.

Diethelm-Okita, B. M., R. Raju, D. K. Okita and B. M. Conti-Fine (1997). Epitope repertoire of human CD4+ T cells on tetanus toxin: identification of immunodominant sequence segments. *J Infect Dis* **175** (2): 382-91.

Donermeyer, D. L. and P. M. Allen (1989). Binding to Ia protects an immunogenic peptide from proteolytic degradation. *J Immunol* **142** (4): 1063-8.

Driessen, C., A. M. Lennon-Dumenil and H. L. Ploegh (2001). Individual cathepsins degrade immune complexes internalized by antigen-presenting cells via Fcgamma receptors. *Eur J Immunol* **31** (5): 1592-601.

Edeling, M. A., C. Smith and D. Owen (2006). Life of a clathrin coat: insights from clathrin and AP structures. *Nat Rev Mol Cell Biol* **7** (1): 32-44.

Engering, A. J., M. Cella, D. Fluitsma, M. Brockhaus, E. C. Hoefsmit, A. Lanzavecchia and J. Pieters (1997). The mannose receptor functions as a high capacity and broad specificity antigen receptor in human dendritic cells. *Eur J Immunol* **27** (9): 2417-25.

Falk, K., O. Rotzschke, S. Stevanovic, G. Jung and H. G. Rammensee (1991). Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* **351** (6324): 290-6.

Falk, K., O. Rotzschke, S. Stevanovic, G. Jung and H. G. Rammensee (1994). Pool sequencing of natural HLA-DR, DQ, and DP ligands reveals detailed peptide motifs, constraints of processing, and general rules. *Immunogenetics* **39** (4): 230-42.

Fenselau, C., M. M. Vestling and R. J. Cotter (1993). Mass spectrometric analysis of proteins. *Curr Opin Biotechnol* **4** (1): 14-9.

Filip, F., G. Fletcher, J. L. Wulff and C. F. Earhart (1973). Solibilization of the cytoplasmic membranc of *Escherichia coli* by the ionic detergent Sodium-Lauryl Sarcosinate. *Journal of Bacteriology* **115** (3): 717-722.

Flyer, D. C., V. Ramakrishna, C. Miller, H. Myers, M. McDaniel, K. Root, C. Flournoy, V. H. Engelhard, D. H. Canaday, J. A. Marto, M. M. Ross, D. F. Hunt, J. Shabanowitz and F. M. White (2002). Identification by mass spectrometry of CD8(+)-T-cell *Mycobacterium tuberculosis* epitopes within the Rv0341 gene product. *Infect Immun* **70** (6): 2926-32.

Flynn, J. L., M. M. Goldstein, K. J. Triebold and B. R. Bloom (1993). Major histocompatibility complex class I-restricted T cells are necessary for protection against M. tuberculosis in mice. *Infect Agents Dis* **2** (4): 259-62.

Flynn, J. L., M. M. Goldstein, K. J. Triebold, B. Koller and B. R. Bloom (1992). Major histocompatibility complex class I-restricted T cells are required for resistance to Mycobacterium tuberculosis infection. *Proc Natl Acad Sci U S A* **89** (24): 12013-7.

Fremont, D. H., M. Matsumura, E. A. Stura, P. A. Peterson and I. A. Wilson (1992). Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb. *Science* **257** (5072): 919-27.

Gagnon, E., S. Duclos, C. Rondeau, E. Chevet, P. H. Cameron, O. Steele-Mortimer, J. Paiement, J. J. Bergeron and M. Desjardins (2002). Endoplasmic reticulum-mediated phagocytosis is a mechanism of entry into macrophages. *Cell* **110** (1): 119-31.

Garin, J., R. Diez, S. Kieffer, J. F. Dermine, S. Duclos, E. Gagnon, R. Sadoul, C. Rondeau and M. Desjardins (2001). The phagosome proteome: insight into phagosome functions. *J Cell Biol.* **152:** 165-80.

Garrett, T. P., M. A. Saper, P. J. Bjorkman, J. L. Strominger and D. C. Wiley (1989). Specificity pockets for the side chains of peptide antigens in HLA-Aw68. *Nature* **342** (6250): 692-6.

Germain, R. N. and D. H. Margulies (1993). The biochemistry and cell biology of antigen processing and presentation. *Annu Rev Immunol* **11**: 403-50.

Gosselin, E. J., K. Wardwell, D. R. Gosselin, N. Alter, J. L. Fisher and P. M. Guyre (1992). Enhanced antigen presentation using human Fc gamma receptor (monocyte/macrophage)-specific immunogens. *J Immunol* **149** (11): 3477-81.

Griffin, J. P., R. Chu and C. V. Harding (1997). Early endosomes and a late endocytic compartment generate different peptide-class II MHC complexes via distinct processing mechanisms. *J Immunol* **158** (4): 1523-32.

Griffiths, E., P. Stevenson, R. Thorpe and H Chart (1985). Naturally occurring antibodies in human sera that react with the iron-regulated outer membrane proteins of *Escherichia coli*. *Infect Immun* **47** (3): 808-13.

Griffiths, E. K., S. Yoonessi and E Neter (1977). Antibody response to enterobacterial lipoprotein of patients with varied infections due to Enterobacteriaceae. *Proc Soc Exp Biol Med* **154** (2): 246-9.

Guagliardi, L. E., B. Koppelman, J. S. Blum, M. S. Marks, P. Cresswell and F. M. Brodsky (1990). Co-localization of molecules involved in antigen processing and presentation in an early endocytic compartment. *Nature* **343**(6254): 133-9.

Gumperz, J. E. and M. B. Brenner (2001). CD1-specific T cells in microbial immunity. *Curr Opin Immunol* **13** (4): 471-8.

Hancock, I. C. and I. R. Poxton (1988). Bacterial cell surface techniques. *John Wiley & Sons*: 279-280.

Hamdan, M. and P. G. Righetti (2002). Modern strategies for protein quantification in proteome analysis: advantages and limitations. *Mass Spectrom Rev* **21** (4): 287-302.

Hammer, J., C. Belunis, D. Bolin, J. Papadopoulos, R. Walsky, J. Higelin, W. Danho, F. Sinigaglia and Z. A. Nagy (1994). High-affinity binding of short peptides to major histocompatibility complex class II molecules by anchor combinations. *Proc Natl Acad Sci U S A* **91** (10): 4456-60.

Hammer, J., E. Bono, F. Gallazzi, C. Belunis, Z. Nagy and F. Sinigaglia (1994). Precise prediction of major histocompatibility complex class II-peptide interaction based on peptide side chain scanning. *J Exp Med* **180** (6): 2353-8.

Haque, M. A., J. W. Hawes and J. S. Blum (2001). Cysteinylation of MHC class II ligands: peptide endocytosis and reduction within APC influences T cell recognition. *J Immunol* **166** (7): 4543-51.

Harding, C. V. (1995). Intracellular organelles involved in antigen processing and the binding of peptides to class II MHC molecules. *Semin Immunol* **7** (6): 355-60.

Harding, C. V. (1996). Class II antigen processing: analysis of compartments and functions. *Crit Rev Immunol* **16** (1): 13-29.

Hellman, J., E. M. Zanzot, P.M. Loiselle, S.F. Amato, K.M. Black, Y. Ge, J.T. Kurnick and HS Warren (1997). Antiserum against *Escherichia coli* J5 contains antibodies reactive with outer membrane proteins of heterologous gram-negative bacteria. *J Infect Dis* **176** (5): 1260-8.

Henriksen, A. Z. and J. A. Maeland (1987). Serum antibodies to outer membrane proteins of *Escherichia coli* in healthy persons and patients with bacteremia. *J Clin Microbiol* **25** (11): 2181-8.

Hewitt, E. W., A. Treumann, N. Morrice, P.J. Tatnell, J. Kay and C Watts (1997). Natural processing sites for human cathepsin E and cathepsin D in tetanus toxin: implications for T cell epitope generation. *J Immunol* **159** (10): 4693-9.

Ho, P. C., D. A. Mutch, K.D. Winkel, A.J. Saul, G.L. Jones, T.J. Doran and C.M. Rzepczyk (1990). Identification of two promiscuous T cell epitopes from tetanus toxin. *Eur J Immunol* **20** (3): 477-83.

Hofstra, H. and J. Dankert (1979). Antigenic cross-reactivity of major outer membrane proteins in enterobacteriaceae species. *J Gen Microbiol* **111** (2): 293-302.

Hofstra, H. and J. Dankert (1980). Major outer membrane proteins: common antigens in enterobacteriaceae species. *J Gen Microbiol* **119** (1): 123-31.

Hofstra, H., J. D. Van Tol, and J. Dankert (1980). Cross-reactivity of major outer membrane proteins of Enterobacteriaceae, studied by crossed immunoelectrophoresis. *J Bacteriol* **143** (1): 328-37.

Hsieh, C. S., P. deRoos, K. Honey, C. Beers and A. Y. Rudensky (2002). A role for cathepsin L and cathepsin S in peptide generation for MHC class II presentation. *J Immunol* **168** (6): 2618-25.

Hunt, D. F., R. A. Henderson, J. Shabanowitz, K. Sakaguchi, H. Michel, N. Sevilir, A. L. Cox, E. Appella and V. H. Engelhard (1992). Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* **255** (5049): 1261-3.

Hunt, D. F., H. Michel, T. A. Dickinson, J. Shabanowitz, A. L. Cox, K. Sakaguchi, E. Appella, H. M. Grey and A. Sette (1992). Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-Ad. *Science* **256** (5065): 1817-20.

Issaq, H. J., T. P. Conrads, G. M. Janini and T. D. Veenstra (2002). Methods for fractionation, separation and profiling of proteins and peptides. *Electrophoresis* **23** (17): 3048-61.

Jensen, P. E. (1993). Acidification and disulfide reduction can be sufficient to allow intact proteins to bind class II MHC. *J Immunol* **150** (8 Pt 1): 3347-56.

Johansen, B. H., F. Vartdal, J. A. Eriksen, E. Thorsby and L. M. Sollid (1996). Identification of a putative motif for binding of peptides to HLA-DQ2. *Int Immunol* **8** (2): 177-82.

Jones, P. P., D. B. Murphy, D. Hewgill and H. O. McDevitt (1979). Detection of a common polypeptide chain in I--A and I--E sub-region immunoprecipitates. *Mol Immunol* **16** (1): 51-60.

Kaliyaperumal, A., M. A. Michaels and S. K. Datta (2002). Naturally processed chromatin peptides reveal a major autoepitope that primes pathogenic T and B cells of lupus. *J Immunol* **168** (5): 2530-7.

Katz, J. F., C. Stebbins, E. Appella and A. J. Sant (1996). Invariant chain and DM edit self-peptide presentation by major histocompatibility complex (MHC) class II molecules. *J Exp Med* **184** (5): 1747-53.

Kerksiek, K. M. and E. G. Pamer (1999). T cell responses to bacterial infection. *Curr Opin Immunol* **11** (4): 400-5.

Krogsgaard, M., Q. J. Li, C. Sumen, J. B. Huppa, M. Huse and M. M. Davis (2005). Agonist/endogenous peptide-MHC heterodimers drive T cell activation and sensitivity. *Nature* **434** (7030): 238-43.

Kropshofer, H., G. J. Hammerling and A. B. Vogt (1997). How HLA-DM edits the MHC class II peptide repertoire: survival of the fittest? *Immunol Today* **18** (2): 77-82.

Kubo, R. T., A. Sette, H. M. Grey, E. Appella, K. Sakaguchi, N. Z. Zhu, D. Arnott, N. Sherman, J. Shabanowitz, H. Michel and et al. (1994). Definition of specific peptide motifs for four major HLA-A alleles. *J Immunol* **152** (8): 3913-24.

Lakey, E. K., L. A. Casten, W. L. Niebling, E. Margoliash and S. K. Pierce (1988). Time dependence of B cell processing and presentation of peptide and native protein antigens. *J Immunol* **140** (10): 3309-14.

Lanzavecchia, A. (1985). Antigen-specific interaction between T and B cells. *Nature* **314** (6011): 537-9.

Lanzavecchia, A. (1990). Receptor-mediated antigen uptake and its effect on antigen presentation to class II-restricted T lymphocytes. *Annu Rev Immunol* **8**: 773-93.

Lanzavecchia, A. (1996). Mechanisms of antigen uptake for presentation. *Curr Opin Immunol* **8** (3): 348-54.

Lauemoller, S. L., C. Kesmir, S. L. Corbet, A. Fomsgaard, A. Holm, M. H. Claesson, S. Brunak and S. Buus (2000). Identifying cytotoxic T cell epitopes from genomic and proteomic information: "The human MHC project." *Rev Immunogenet* **2** (4): 477-91.

Lee, P., G. R. Matsueda and P. M. Allen (1988). T cell recognition of fibrinogen. A determinant on the A alpha-chain does not require processing. *J Immunol* **140** (4): 1063-8.

Lemmel, C. and S. Stevanovic (2003). The use of HPLC-MS in T-cell epitope identification. *Methods* **29** (3): 248-59.

Lemmel, C., S. Weik, U. Eberle, J. Dengjel, T. Kratt, H. D. Becker, H. G. Rammensee and S. Stevanovic (2004). Differential quantitative analysis of MHC ligands by mass spectrometry using stable isotope labeling. *Nat Biotechnol* **22** (4): 450-4.

Li, P., M. A. Haque and J. S. Blum (2002). Role of disulfide bonds in regulating antigen processing and epitope selection. *J Immunol* **169** (5): 2444-50.

Lich, J. D., J. F. Elliott and J. S. Blum (2000). Cytoplasmic processing is a prerequisite for presentation of an endogenous antigen by major histocompatibility complex class II proteins. *J Exp Med* **191** (9): 1513-24.

Liljedahl, M., T. Kuwana, W. P. Fung-Leung, M. R. Jackson, P. A. Peterson and L. Karlsson (1996). HLA-DO is a lysosomal resident which requires association with HLA-DM for efficient intracellular transport. *Embo J* **15** (18): 4817-24.

Lotteau, V., L. Teyton, A. Peleraux, T. Nilsson, L. Karlsson, S. L. Schmid, V. Quaranta and P. A. Peterson (1990). Intracellular transport of class II MHC molecules directed by invariant chain. *Nature* **348** (6302): 600-5.

Lowry, O. H., N. J. Rosebrough, A. L. Farr and R. L. Randall (1951). Protein measurement with the Folin phenol reagent. *Journal of Biological Chemistry* **193**: 265-275.

Maehr, R., H. C. Hang, J. D. Mintern, Y. M. Kim, A. Cuvillier, M. Nishimura, K. Yamada, K. Shirahama-Noda, I. Hara-Nishimura and H. L. Ploegh (2005). Asparagine endopeptidase is not essential for class II MHC antigen presentation but is required for processing of cathepsin L in mice. *J Immunol* **174** (11): 7066-74.

Manoury, B., E. W. Hewitt, N. Morrice, P. M. Dando, A. J. Barrett and C. Watts (1998). An asparaginyl endopeptidase processes a microbial antigen for class II MHC presentation. *Nature* **396** (6712): 695-9.

Manoury, B., D. Mazzeo, L. Fugger, N. Viner, M. Ponsford, H. Streeter, G. Mazza, D. C. Wraith and C. Watts (2002). Destructive processing by asparagine endopeptidase limits presentation of a dominant T cell epitope in MBP. *Nat Immunol* **3** (2): 169-74.

Manoury, B., D. Mazzeo, D. N. Li, J. Billson, K. Loak, P. Benaroch and C. Watts (2003). Asparagine endopeptidase can initiate the removal of the MHC class II invariant chain chaperone. *Immunity* **18** (4): 489-98.

Maric, M., B. Arunachalam, U. T. Phan, C. Dong, W. S. Garrett, K. S. Cannon, C. Alfonso, L. Karlsson, R. A. Flavell and P. Cresswell (2001). Defective antigen processing in GILT-free mice. *Science* **294** (5545): 1361-5.

Matsumura, M., D. H. Fremont, P. A. Peterson and I. A. Wilson (1992). Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science* **257** (5072): 927-34.

McCoy, K. L., M. Noone, J. K. Inman and R. Stutzman (1993). Exogenous antigens internalized through transferrin receptors activate CD4+ T cells. *J Immunol* **150** (5): 1691-701.

Meiring, H. D., B. Kuipers, J. A. van Gaans-van den Brink, M. C. Poelen, H. Timmermans, G. Baart, H. Brugghe, J. van Schie, C. J. Boog, A. P. de Jong and C. A. van Els (2005). Mass tag-assisted identification of naturally processed HLA class II-presented meningococcal peptides recognized by CD4+ T lymphocytes. *J Immunol* **174** (9): 5636-43.

Michrom Bioresources, I. An automated 2D Capillary HPLC System Coupled to Nanoflow ESI-MS/MS for high sensitivity , High Resolution Peptide Mapping and Protein Identification of Complex Proteomics Samples. **Application Note No. 111**.

Minino, A. M., E. Arias, K.D. Kockanek, S.L. Murphy and B.L. Smith (2002). Deaths: final data for 2000. *Natl Vital Stat Rep* ***50*** *Rep*50 (15): 1-119.

Molloy, M. P., B. R. Herbert, M. B. Slade, T. Rabilloud, A. S. Nouwens, K. L. Williams and A. A. Gooley (2000). Proteomic analysis of the *Escherichia coli* outer membrane. *Eur J Biochem* **267** (10): 2871-81.

Moudgil, K. D., E. E. Sercarz and I. S. Grewal (1998). Modulation of the immunogenicity of antigenic determinants by their flanking residues. *Immunol Today* **19** (5): 217-20.

Moss, C. X., S. P. Matthews, D. J. Lamont and C. Watts (2005). Asparagine deamidation perturbs antigen presentation on class II major histocompatibility complex molecules. *J Biol Chem* **280** (18): 18498-503.

Musson, J. A., R. D. Hayward, A. A. Delvig, C. E. Hormaeche, V. Koronakis and J. H. Robinson (2002). Processing of viable Salmonella typhimurium for presentation of a CD4 T cell epitope from the Salmonella invasion protein C (SipC). *Eur J Immunol* **32** (9): 2664-71.

Musson, J. A., N. Walker, H. Flick-Smith, E. D. Williamson and J. H. Robinson (2003). Differential processing of CD4 T-cell epitopes from the protective antigen of Bacillus anthracis. *J Biol Chem* **278** (52): 52425-31.

Nakagawa, T., W. Roth, P. Wong, A. Nelson, A. Farr, J. Deussing, J. A. Villadangos, H. Ploegh, C. Peters and A. Y. Rudensky (1998). Cathepsin L: critical role in Ii degradation and CD4 T cell selection in the thymus. *Science* **280** (5362): 450-3.

Nakagawa, T. Y., W. H. Brissette, P. D. Lira, R. J. Griffiths, N. Petrushova, J. Stock, J. D. McNeish, S. E. Eastman, E. D. Howard, S. R. Clarke, E. F. Rosloniec, E. A. Elliott and A. Y. Rudensky (1999). Impaired invariant chain degradation and antigen presentation and diminished collagen-induced arthritis in cathepsin S null mice. *Immunity* **10** (2): 207-17.

Nagele, E., M. Vollmer, and P. Horth (2004). Improved 2D nano-LC/MS for proteomics applications: a comparative analysis using yeast proteome. *J Biomol Tech* **15** (2): 134-43.

Nelson, C. A., R. W. Roof, D. W. McCourt and E. R. Unanue (1992). Identification of the naturally processed form of hen egg white lysozyme bound to the murine major histocompatibility complex class II molecule I-Ak. *Proc Natl Acad Sci U S A* **89** (16): 7380-3.

Nepom, G. T., J. D. Lippolis, F. M. White, S. Masewicz, J. A. Marto, A. Herman, C. J. Luckey, B. Falk, J. Shabanowitz, D. F. Hunt, V. H. Engelhard and B. S. Nepom

(2001). Identification and modulation of a naturally processed T cell epitope from the diabetes-associated autoantigen human glutamic acid decarboxylase 65 (hGAD65). *Proc Natl Acad Sci U S A* **98** (4): 1763-8.

Newcomb, J. R. and P. Cresswell (1993). Characterization of endogenous peptides bound to purified HLA-DR molecules and their absence from invariant chain-associated alpha beta dimers. *J Immunol* **150** (2): 499-507.

Nicolle, L. E., E. Ujack, et al. (1988). Immunoblot analysis of serologic response to outer membrane proteins of *Escherichia coli* in elderly individuals with urinary tract infections. *J Clin Microbiol* **26** (10): 2087-91.

Nilsson, C. L. (2002). Bacterial proteomics and vaccine development. *Am J Pharmacogenomics* **2** (1): 59-65.

O'Sullivan, D., J. Sidney, E. Appella, L. Walker, L. Phillips, S. M. Colon, C. Miles, R. W. Chesnut and A. Sette (1990). Characterization of the specificity of peptide binding to four DR haplotypes. *J Immunol* **145** (6): 1799-808.

O'Sullivan, D., T. Arrhenius, J. Sidney, M. F. Del Guercio, M. Albertson, M. Wall, C. Oseroff, S. Southwood, S. M. Colon, F. C. Gaeta and et al. (1991). On the interaction of promiscuous antigenic peptides with different DR alleles. Identification of common structural motifs. *J Immunol* **147** (8): 2663-9.

Overbeeke, N. and B. Lugtenberg (1980). Major outer membrane proteins of *Escherichia coli* strains of human origin. *J Gen Microbiol* **121** (2): 373-80.

Pamer, E. and P. Cresswell (1998). Mechanisms of MHC class I--restricted antigen processing. *Annu Rev Immunol* **16**: 323-58.

Pamer, E. G. (1994). Direct sequence identification and kinetic analysis of an MHC class I-restricted Listeria monocytogenes CTL epitope. *J Immunol* **152** (2): 686-94.

Panina-Bordignon, P., A. Tan, A. Termijtelen, S. Demotz, G. Corradin and A. Lanzavecchia (1989). Universally immunogenic T cell epitopes: promiscuous binding to human MHC class II and promiscuous recognition by T cells. *Eur J Immunol* **19** (12): 2237-42.

Pappin, D. H., H. P. and B. A.J. (1993). Rapid identification of proteins by peptide mass-fingerprinting. *Curr Biol.* **3** (6): 327-32.

Parker, K. C., M. A. Bednarek and J. E. Coligan (1994). Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *J Immunol* **152** (1): 163-75.

Peakman, M., E. J. Stevens, T. Lohmann, P. Narendran, J. Dromey, A. Alexander, A. J. Tomlinson, M. Trucco, J. C. Gorga and R. M. Chicz (1999). Naturally processed and presented epitopes of the islet cell autoantigen IA-2 eluted from HLA-DR4. *J Clin Invest* **104** (10): 1449-57.

Peters, P. J., J. J. Neefjes, V. Oorschot, H. L. Ploegh and H. J. Geuze (1991). Segregation of MHC class II molecules from MHC class I molecules in the Golgi complex for transport to lysosomal compartments. *Nature* **349** (6311): 669-76.

Phelps, R. G., A. N. Turner and A. J. Rees (1996). Direct identification of naturally processed autoantigen-derived peptides bound to HLA-DR15. *J Biol Chem* **271** (31): 18549-53.

Pierre, P. and I. Mellman (1998). Exploring the mechanisms of antigen processing by cell fractionation. *Curr Opin Immunol* **10** (2): 145-53.

Pieters, J. (2000). MHC class II-restricted antigen processing and presentation. *Adv Immunol* **75**: 159-208.

Pieters, J., O. Bakke and B. Dobberstein (1993). The MHC class II-associated invariant chain contains two endosomal targeting signals within its cytoplasmic tail. *J Cell Sci* **106 (Pt 3)**: 831-46.

Pieters, J., H. Horstmann, O. Bakke, G. Griffiths and J. Lipp (1991). Intracellular transport and localization of major histocompatibility complex class II molecules and associated invariant chain. *J Cell Biol* **115** (5): 1213-23.

Pinet, V., M. S. Malnati and E. O. Long (1994). Two processing pathways for the MHC class II-restricted presentation of exogenous influenza virus antigen. *J Immunol* **152** (10): 4852-60.

Pinet, V. M. and E. O. Long (1998). Peptide loading onto recycling HLA-DR molecules occurs in early endosomes. *Eur J Immunol* **28** (3): 799-804.

Pluger, E. B., M. Boes, C. Alfonso, C. J. Schroter, H. Kalbacher, H. L. Ploegh and C. Driessen (2002). Specific role for cathepsin S in the generation of antigenic peptides in vivo. *Eur J Immunol* **32** (2): 467-76.

Porcelli, S. A. (1995). The CD1 family: a third lineage of antigen-presenting molecules. *Adv Immunol* **59**: 1-98.

Schaegger, H. and G. vonJagow (1987). Tricine-Sodium dodecyl sulfate Polyacrylamide Gel Electrophoresis for the Separation of Proteins in the Range from 1 to 100 kDa. *Anal Biochem.* **166**: 368-379.

Raju, R., B. Diethelm-Okita, et al. (1996). "Epitope repertoire of human CD4+ lines propagated with tetanus toxoid or with synthetic tetanus toxin sequences." *J Autoimmun* **9** (1): 79-88.

Ramachandra, L., R. S. Chu, D. Askew, E. H. Noss, D. H. Canaday, N. S. Potter, A. Johnsen, A. M. Krieg, J. G. Nedrud, W. H. Boom and C. V. Harding (1999). Phagocytic antigen processing and effects of microbial products on antigen processing and T-cell responses. *Immunol Rev* **168**: 217-39.

Ramachandra, L. and C. V. Harding (2000). Phagosomes acquire nascent and recycling class II MHC molecules but primarily use nascent molecules in phagocytic antigen processing. *J Immunol* **164** (10): 5103-12.

Ramachandra, L., E. Noss, W. H. Boom and C. V. Harding (1999). Phagocytic processing of antigens for presentation by class II major histocompatibility complex molecules. *Cell Microbiol* **1** (3): 205-14.

Rammensee, H. G., K. Falk and O. Rotzschke (1993). MHC molecules as peptide receptors. *Curr Opin Immunol* **5** (1): 35-44.

Rammensee, H. G., K. Falk and O. Rotzschke (1993). Peptides naturally presented by MHC class I molecules. *Annu Rev Immunol* **11**: 213-44.

Rammensee, H. G. (1995). Chemistry of peptides associated with MHC class I and class II molecules. *Curr Opin Immunol* **7** (1): 85-96.

Reece, J. C., H. M. Geysen, and S.J. Rodda (1993). "Mapping the major human T helper epitopes of tetanus toxin. The emerging picture." *J Immunol* **151**(11): 6175-84.

Riberdy, J. M., J. R. Newcomb, M. J. Surman, J. A. Barbosa and P. Cresswell (1992). HLA-DR molecules from an antigen-processing mutant cell line are associated with invariant chain peptides. *Nature* **360** (6403): 474-7.

Riese, R. J. and H. A. Chapman (2000). Cathepsins and compartmentalization in antigen presentation. *Curr Opin Immunol* **12** (1): 107-13.

Riese, R. J., P. R. Wolf, D. Bromme, L. R. Natkin, J. A. Villadangos, H. L. Ploegh and H. A. Chapman (1996). Essential role for cathepsin S in MHC class II-associated invariant chain processing and peptide loading. *Immunity* **4** (4): 357-66.

Ringrose, J. H., H. D. Meiring, D. Speijer, T. E. Feltkamp, C. A. van Els, A. P. de Jong and J. Dankert (2004). Major histocompatibility complex class I peptide presentation after Salmonella enterica serovar typhimurium infection assessed via stable isotope tagging of the B27-presented peptide repertoire. *Infect Immun* **72** (9): 5097-105.

Robinson, J., M. J. Waller, P. Parham, N. de Groot, R. Bontrop, L. J. Kennedy, P. Stoehr and S. G. Marsh (2003). IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Res* **31** (1): 311-4.

Robinson, J. H. and A. A. Delvig (2002). Diversity in MHC class II antigen presentation. *Immunology* **105** (3): 252-62.

Roche, P. A. and P. Cresswell (1990). High-affinity binding of an influenza hemagglutinin-derived peptide to purified HLA-DR. *J Immunol* **144** (5): 1849-56.

Roche, P. A. and P. Cresswell (1990). Invariant chain association with HLA-DR molecules inhibits immunogenic peptide binding. *Nature* **345** (6276): 615-8.

Roche, P. A., C. L. Teletski, E. Stang, O. Bakke and E. O. Long (1993). Cell surface HLA-DR-invariant chain complexes are targeted to endosomes by rapid internalization. *Proc Natl Acad Sci U S A* **90** (18): 8581-5.

Rock, K. L., B. Benacerraf and A. K. Abbas (1984). Antigen presentation by hapten-specific B lymphocytes. I. Role of surface immunoglobulin receptors. *J Exp Med* **160** (4): 1102-13.

Rocque, W. J., R. T. Coughlin, and E.J. McGroarty (1987). Lipopolysaccharide tightly bound to porin monomers and trimers from *Escherichia coli* K-12. *J Bacteriol* 169 (9): 4003-10.

Rosat, J. P., E. P. Grant, E. M. Beckman, C. C. Dascher, P. A. Sieling, D. Frederique, R. L. Modlin, S. A. Porcelli, S. T. Furlong and M. B. Brenner (1999). CD1-restricted microbial lipid antigen-specific recognition found in the CD8+ alpha beta T cell pool. *J Immunol* **162** (1): 366-71.

Rotzschke, O., K. Falk, K. Deres, H. Schild, M. Norda, J. Metzger, G. Jung and H. G. Rammensee (1990). Isolation and analysis of naturally processed viral peptides as recognized by cytotoxic T cells. *Nature* **348** (6298): 252-4.

Rudensky, A., P. Preston-Hurlburt, S. C. Hong, A. Barlow and C. A. Janeway, Jr. (1991). Sequence analysis of peptides bound to MHC class II molecules. *Nature* **353** (6345): 622-7.

Ruppert, J., J. Sidney, E. Celis, R. T. Kubo, H. M. Grey and A. Sette (1993). Prominent role of secondary anchor residues in peptide binding to HLA-A2.1 molecules. *Cell* **74** (5): 929 37.

Russo, T. A. and J. R. Johnson (2003). Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes Infect* **5** (5): 449-56.

Russo, T. A., C. D. McFadden, U.B. Carlino-Macdonald, J.M. Beanan, R. Olson and G.E. Wilding (2003). The Siderophore receptor IroN of extraintestinal pathogenic *Escherichia coli* is a potential vaccine candidate. *Infect Immun* **71** (12): 7164-9.

Sallusto, F., M. Cella, C. Danieli and A. Lanzavecchia (1995). Dendritic cells use macropinocytosis and the mannose receptor to concentrate macromolecules in the major histocompatibility complex class II compartment: downregulation by cytokines and bacterial products. *J Exp Med* **182** (2): 389-400.

Sanderson, S., D. J. Campbell and N. Shastri (1995). Identification of a CD4+ T cell-stimulating antigen of pathogenic bacteria by expression cloning. *J Exp Med* **182** (6): 1751-7.

Sarangarajan, R. and S. P. Apte (2005). Melanization and phagocytosis: implications for age related macular degeneration. *Mol Vis* **11**: 482-90

Schaegger, H. and G. vonJagow (1987). Tricine-Sodium dodecyl sulfate-Polyacrylamide Gel Electrophoresis for the Separation of Proteins in the Range from 1 to 100 kDa. *Anal Biochem.* **166**: 368-379.

Schirle, M., T. Weinschenk and S. Stevanovic (2001). Combining computer algorithms with experimental approaches permits the rapid and accurate identification of T cell epitopes from defined antigens. *J Immunol Methods* **257** (1-2): 1-16.

Schneider, S. C., J. Ohmen, L. Fosdick, B. Gladstone, J. Guo, A. Ametani, E. E. Sercarz and H. Deng (2000). Cutting edge: introduction of an endopeptidase cleavage motif into a determinant flanking region of hen egg lysozyme results in enhanced T cell determinant display. *J Immunol* **165** (1): 20-3.

Schumacher, T. N., M. L. De Bruijn, L. N. Vernie, W. M. Kast, C. J. Melief, J. J. Neefjes and H. L. Ploegh (1991). Peptide selection by MHC class I molecules. *Nature* **350** (6320): 703-6.

Schreuder, G. M., C. K. Hurley, S. G. Marsh, M. Lau, M. Fernandez-Vina, H. J. Noreen, M. Setterholm and M. Maiers (2005). The HLA Dictionary 2004: a summary of HLA-A, -B, -C, -DRB1/3/4/5 and -DQB1 alleles and their association with serologically defined HLA-A, -B, -C, -DR and -DQ antigens. *Int J Immunogenet* **32** (1): 19-69.

Schroter, C. J., M. Braun, J. Englert, H. Schmid and H. Kalbacher (1999). A rapid method to separate endosomes from lysosomal contents using differential centrifugation and hypotonic lysis of lysosomes. *J Immunol Methods* **227** (1-2): 161-8.

Seidah, N. G. and M. Chretien (1997). Eukaryotic protein processing: endoproteolysis of precursor proteins. *Curr Opin Biotechnol* **8** (5): 602-7.

Sercarz, E. E., P. V. Lehmann, A. Ametani, G. Benichou, A. Miller and K. Moudgil (1993). Dominance and crypticity of T cell antigenic determinants. *Annu Rev Immunol* **11**: 729-66.

Sercarz, E. E. and E. Maverakis (2003). Mhc-guided processing: binding of large antigen fragments. *Nat Rev Immunol* **3** (8): 621-9.

Sette, A., L. Adorini, S. M. Colon, S. Buus and H. M. Grey (1989). Capacity of intact proteins to bind to MHC class II molecules. *J Immunol* **143** (4): 1265-7.

Sherman, M. A., D. A. Weber and P. E. Jensen (1995). DM enhances peptide binding to class II MHC by release of invariant chain-derived peptide. *Immunity* **3** (2): 197-205.

Shi, G. P., J. A. Villadangos, G. Dranoff, C. Small, L. Gu, K. J. Haley, R. Riese, H. L. Ploegh and H. A. Chapman (1999). Cathepsin S required for normal MHC class II peptide loading and germinal center development. *Immunity* **10** (2): 197-206.

Sijts, A. J., A. Neisig, J. Neefjes and E. G. Pamer (1996). Two *Listeria monocytogenes* CTL epitopes are processed from the same antigen with different efficiencies. *J Immunol* **156** (2): 683-92.

Silverstein, S. C. (1995). Phagocytosis of microbes: insights and prospects. *Trends Cell Biol* **5** (3): 141-2.

Simitsek, P. D., D. G. Campbell, A. Lanzavecchia, N. Fairweather and C. Watts (1995). Modulation of antigen processing by bound antibodies can boost or suppress class II major histocompatibility complex presentation of different T cell determinants. *J Exp Med* **181** (6): 1957-63.

Sinnathamby, G., M. Maric, P. Cresswell and L. C. Eisenlohr (2004). Differential requirements for endosomal reduction in the presentation of two H2-E(d)-restricted epitopes from influenza hemagglutinin. *J Immunol* **172** (11): 6607-14.

Sloan, V. S., P. Cameron, G. Porter, M. Gammon, M. Amaya, E. Mellins and D. M. Zaller (1995). Mediation by HLA-DM of dissociation of peptides from HLA-DR. *Nature* **375** (6534): 802-6.

Snider, D. P. and D. M. Segal (1989). Efficiency of antigen presentation after antigen targeting to surface IgD, IgM, MHC, Fc gamma RII, and B220 molecules on murine splenic B cells. *J Immunol* **143** (1): 59-65.

Sommerer, N., D. Centeno and M. Rossignol (2006). Peptide mass fingerprinting: identification of proteins by maldi-tof. *Methods Mol Biol* **355**: 219-34.

Stahl, P., P. H. Schlesinger, E. Sigardson, J. S. Rodman and Y. C. Lee (1980). Receptor-mediated pinocytosis of mannose glycoconjugates by macrophages: characterization and evidence for receptor recycling. *Cell* **19** (1): 207-15.

Steinman, R. M., I. S. Mellman, W. A. Muller and Z. A. Cohn (1983). Endocytosis and the recycling of plasma membrane. *J Cell Biol* **96** (1): 1-27.

Stenger, S., R. J. Mazzaccaro, K. Uyemura, S. Cho, P. F. Barnes, J. P. Rosat, A. Sette, M. B. Brenner, S. A. Porcelli, B. R. Bloom and R. L. Modlin (1997). Differential effects of cytolytic T cell subsets on intracellular infection. *Science* **276** (5319): 1684-7.

Stenger, S. and R. L. Modlin (1998). Cytotoxic T cell responses to intracellular pathogens. *Curr Opin Immunol* **10** (4): 471-7.

Stern, L. J., J. H. Brown, T. S. Jardetzky, J. C. Gorga, R. G. Urban, J. L. Strominger and D. C. Wiley (1994). Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* **368** (6468): 215-21.

Strauss, O. (2005). The retinal pigment epithelium in visual function. *Physiol Rev* **85** (3): 845-81.

Sugita, M., M. Cernadas and M. B. Brenner (2004). New insights into pathways for CD1-mediated antigen presentation. *Curr Opin Immunol* **16** (1): 90-5.

Swanson, J. A. and C. Watts (1995). Macropinocytosis. *Trends Cell Biol* **5** (11): 424-8.

Tao, W. A. and R. Aebersold (2003). Advances in quantitative proteomics via stable isotope tagging and mass spectrometry. *Curr Opin Biotechnol* **14** (1): 110-8.

Tsai, C. and C. Frasch (1982). "A sensitive silver stain for detecting LPS in polyacrylamide gels." *Analytical biochemistry* **119**: 115-119.

Teyton, L., D. O'Sullivan, P. W. Dickson, V. Lotteau, A. Sette, P. Fink and P. A. Peterson (1990). Invariant chain distinguishes between the exogenous and endogenous antigen presentation pathways. *Nature* **348** (6296): 39-44.

Townsend, A. and H. Bodmer (1989). Antigen recognition by class I-restricted T lymphocytes. *Annu Rev Immunol* **7**: 601-24.

Trombetta, E. S. and I. Mellman (2005). Cell biology of antigen processing in vitro and in vivo. *Annu Rev Immunol* **23**: 975-1028.

Turk, B., D. Turk and V. Turk (2000). Lysosomal cysteine proteases: more than scavengers. *Biochim Biophys Acta* **1477** (1-2): 98-111.

Unanue, E. R. (1984). Antigen-presenting function of the macrophage. *Annu Rev Immunol* **2**: 395-428.

Van Bleek, G. M. and S. G. Nathenson (1990). Isolation of an endogenously processed immunodominant viral peptide from the class I H-2Kb molecule. *Nature* **348** (6298): 213-6.

Vignali, D. A., R. G. Urban, R. M. Chicz and J. L. Strominger (1993). Minute quantities of a single immunodominant foreign epitope are presented as large nested sets by major histocompatibility complex class II molecules. *Eur J Immunol* **23** (7): 1602-7.

Villadangos, J. A. and H. L. Ploegh (2000). Proteolysis in MHC class II antigen presentation: who's in charge? *Immunity* **12** (3): 233-9.

Villadangos, J. A., R. J. Riese, C. Peters, H. A. Chapman and H. L. Ploegh (1997). Degradation of mouse invariant chain: roles of cathepsins S and D and the influence of major histocompatibility complex polymorphism. *J Exp Med* **186** (4): 549-60.

Walden, P. (1996). T-cell epitope determination. *Curr Opin Immunol* **8** (1): 68-74.

Wang, H. and S. Hanash (2003). Multi-dimensional liquid phase based separations in proteomics. *J Chromatogr B Analyt Technol Biomed Life Sci* **787** (1): 11-8.

Watts, C. (1997). Capture and processing of exogenous antigens for presentation on MHC molecules. *Annu Rev Immunol* **15**: 821-50.

Watts, C. (2001). Antigen processing in the endocytic compartment. *Curr Opin Immunol* **13** (1): 26-31.

Watts, C. (2001). Immunology. Antigen presentation--losing its shine in the absence of GILT. *Science* **294** (5545): 1294-5.

Watts, C. (2002). Phagocytosis: how the phagosome became the phag-ER-some. *Curr Biol* **12:** R666-8.

Watts, C. (2004). The exogenous pathway for antigen presentation on major histocompatibility complex class II and CD1 molecules. *Nat Immunol* **5**(7): 685-92.

Watts, C., A. Antoniou, B. Manoury, E. W. Hewitt, L. M. McKay, L. Grayson, N. F. Fairweather, P. Emsley, N. Isaacs and P. D. Simitsek (1998). Modulation by epitope-specific antibodies of class II MHC-restricted presentation of the tetanus toxin antigen. *Immunol Rev* **164**: 11-6.

Watts, C., C. X. Moss, D. Mazzeo, M. A. West, S. P. Matthews, D. N. Li and B. Manoury (2003). Creation versus destruction of T cell epitopes in the class II MHC pathway. *Ann N Y Acad Sci* **987**: 9-14.

Weber, D. A., B. D. Evavold and P. E. Jensen (1996). Enhanced dissociation of HLA-DR-bound peptides in the presence of HLA-DM. *Science* **274** (5287): 618-20.

West, M. A., A. R. Prescott, E. L. Eskelinen, A. J. Ridley and C. Watts (2000). Rac is required for constitutive macropinocytosis by dendritic cells but does not control its downregulation. *Curr Biol* **10** (14): 839-48.

Williams, K. M. and E. C. Bigley, 3rd (2004). Identification of an I-Ed-restricted T-cell epitope of *Escherichia coli* outer membrane protein F. *Infect Immun* **72**(7): 3907-13.

Williams, K. M., E. C. Bigley, 3rd, et al. (2000). Identification of murine B-cell and T-cell epitopes of *Escherichia coli* outer membrane protein F with synthetic polypeptides. *Infect Immun* **68** (5): 2535-45.

Wilson, R. A., R. S. Curwen, S. Braschi, S. L. Hall, P. S. Coulson and P. D. Ashton (2004). From genomes to vaccines via the proteome. *Mem Inst Oswaldo Cruz* **99** (5 Suppl 1): 45-50.

Xiang, S. D., A. Scholzen, G. Minigo, C. David, V. Apostolopoulos, P. L. Mottram and M. Plebanski (2006). Pathogen recognition and development of particulate vaccines: does size matter? *Methods* **40** (1): 1-9.

Zeligs, J. D. and S. H. Wollman (1977). Ultrastructure of blebbing and phagocytosis of blebs by hyperplastic thyroid epithelial cells in vivo. *J Cell Biol* **72** (3): 584-94.

Zhang, T., Y. Maekawa, J. Hanba, T. Dainichi, B. F. Nashed, H. Hisaeda, T. Sakai, T. Asao, K. Himeno, R. A. Good and N. Katunuma (2000). Lysosomal cathepsin B plays an important role in antigen processing, while cathepsin D is involved in degradation of the invariant chain inovalbumin-immunized mice. *Immunology* **100** (1): 13-20.

Zinkernagel, R. M. and P. C. Doherty (1997). The discovery of MHC restriction. *Immunol Today* **18** (1): 14-7.

Zou, J., A.N. Turner and R.G. Phelps (2003). Mass and composition matrix-assisted laser desorption ionisation time of flight analysis enabling inference of the sequence of most peptides where the protein of origin is known. *Anal Chem* **75** (11): 2653-62

Zou, J., A. N. Turner and R. G. Phelps (2004). Trace labeling of proteins with stable isotopes to identify fragments in complex mixtures. *Anal Chem* **76** (5): 1445-52.